



UNIVERSIDADE ESTADUAL DE CAMPINAS

Instituto de Física “Gleb Wataghin”

YUGO MAFRA KUNO

ELABORAÇÃO DE UM MÉTODO GLOBAL DE RECONSTRUÇÃO DE CHUVEIROS
EXTENSOS UTILIZANDO TEORIA DA INFORMAÇÃO E OTIMIZAÇÃO
MATEMÁTICA

CAMPINAS

2017

YUGO MAFRA KUNO

ELABORAÇÃO DE UM MÉTODO GLOBAL DE RECONSTRUÇÃO DE CHUVEIROS
EXTENSOS UTILIZANDO TEORIA DA INFORMAÇÃO E OTIMIZAÇÃO
MATEMÁTICA

Dissertação apresentada ao Instituto de Física “Gleb Wataghin” da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em Física.

Orientador: Prof. Dr. José Augusto Chinellato

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL
DA DISSERTAÇÃO DEFENDIDA PELO ALUNO YUGO
MAFRA KUNO E ORIENTADA PELO PROF. DR. JOSÉ
AUGUSTO CHINELLATO.

CAMPINAS

2017

Agência(s) de fomento e nº(s) de processo(s): CAPES

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Física Gleb Wataghin
Lucimeire de Oliveira Silva da Rocha - CRB 8/9174

K963e Kuno, Yugo Mafra, 1982-
Elaboração de um método global de reconstrução de chuviros extensos utilizando teoria da informação e otimização matemática / Yugo Mafra Kuno. – Campinas, SP : [s.n.], 2017.

Orientador: José Augusto Chinellato.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Física Gleb Wataghin.

1. Raios cósmicos. 2. Chuviros de raios cósmicos. 3. Calorimetria. 4. Teoria da informação. 5. Otimização matemática. I. Chinellato, José Augusto, 1950-. II. Universidade Estadual de Campinas. Instituto de Física Gleb Wataghin. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Elaboration of a global method for the reconstruction of extensive air showers utilizing information theory and mathematical optimization

Palavras-chave em inglês:

Cosmic rays

Cosmic ray showers

Calorimetry

Information theory

Mathematical optimization

Área de concentração: Física

Titulação: Mestre em Física

Banca examinadora:

José Augusto Chinellato [Orientador]

Edmilson José Tonelli Manganote

Fernando Catalani

Data de defesa: 28-08-2017

Programa de Pós-Graduação: Física

MEMBROS DA COMISSÃO JULGADORA DA DISSERTAÇÃO DE MESTRADO DE **YUGO MAFRA KUNO - RA 033324** APRESENTADA E APROVADA AO INSTITUTO DE FÍSICA “GLEB WATAGHIN”, DA UNIVERSIDADE ESTADUAL DE CAMPINAS, EM 28 / 08 / 2017.

COMISSÃO JULGADORA:

- **Prof. Dr. José Augusto Chinellato – Orientador – DRCC/IFGW/UNICAMP**
- **Prof. Dr. Fernando Catalani – EEL/USP**
- **Prof. Dr. Edmilson José Tonelli Manganote – DRCC/IFGW/UNICAMP**

OBS.: Informo que as assinaturas dos respectivos professores membros da banca constam na ata de defesa já juntada no processo vida acadêmica do aluno.

AGRADECIMENTOS

Agradeço imensamente ao Prof. Dr. José Augusto Chinellato por ter aceitado me orientar, pelo apoio, por tudo que ele me ensinou, pelas conversas, e por toda sua paciência ao longo deste percurso. Serei sempre muito grato pela oportunidade de trabalhar sob a sua supervisão.

Agradeço também a todos os outros que estiveram ao meu lado, pelo incentivo para que fosse possível chegar até aqui. Seu apoio foi inestimável, em todos os momentos.

RESUMO

O fluxo de raios cósmicos de ultra-alta energia (superiores a $10^{18}eV$) que chega ao topo da atmosfera é baixo, tal que torna-se necessário estudar em detalhes os chuveiros atmosféricos extensos de modo a se determinar a composição e a energia da partícula primária. Dispõe-se de algoritmos que simulam processos de *QED* e *QCD*, e descrevem a difusão das partículas na atmosfera, fornecendo o estado $(\vec{x}, \vec{p}, id, t)$ de cada uma. Entretanto, é também necessário parametrizar as funções que descrevem sua distribuição ao longo da cascata atmosférica. O *método da entropia máxima* (MEM) [1, 2] permite que se parametrize a distribuição de partículas em função dos momentos estatísticos desses perfis, que são menos afetados pelo ruído. A maximização da entropia também permite que a distribuição seja inferida minimizando o enviesamento que ocorre, por exemplo, no ajuste de curvas tradicionalmente realizado. No presente trabalho foi estudada aplicação do MEM na análise do perfil longitudinal das componentes eletrônica e muônica geradas por simulação no CORSIKA [3], já conhecendo as características da partícula primária, de forma a avaliar os benefícios dessa metodologia no estudo de chuveiros atmosféricos.

ABSTRACT

The flux of ultra-high energy cosmic rays (above $10^{18}eV$) that reach the upper atmosphere is low, so that it is necessary to thoroughly study air showers in order to determine the composition and energy of the primary particle. There are algorithms that simulate *QED* and *QCD* processes, and track the particle diffusion in the atmosphere, providing the state $(\vec{x}, \vec{p}, id, t)$ of each one. Still, it is also necessary to parameterize the functions that describe its distribution throughout the atmospheric cascade. The *maximum entropy method* (MEM) [1, 2] allows the parameterization of the particle distribution with the statistical moments as variables, which are less affected by noise. The maximum entropy also allows the inference of the distribution minimizing the bias, which is high in inference processes such as the traditional curve fitting. The application of MEM in the analysis of the longitudinal profile of the electronic and muonic components generated by simulation with CORSIKA [3] was the object of study in this research work, having as a prior the characteristics of the primary particle, in order to evaluate the benefits of the method in the study of air showers.

LISTA DE FIGURAS

2.1	Espectro normalizado de energia de raios cósmicos. [10]	16
2.2	Espectro normalizado de energia de raios cósmicos na faixa de ultra-altas energias [12]	17
2.3	Representação do desenvolvimento de um chuveiro atmosférico, com o logaritmo do número de partículas em função da profundidade na atmosfera à esquerda e as interações típicas do primário e dos secundários à direita. [16]	20
2.4	Desenvolvimento de cascata a 3027m de altitude em uma câmara de nuvens que contém placas de chumbo de 1.3cm de espessura (faixa escuras numeradas). A área da seção de choque da câmara era de $0.5 \times 0.3m^2$. [17]	21
2.5	Perfis longitudinais de número de partículas carregadas para primários de próton (em linhas sólidas pretas) e ferro (em linhas tracejadas vermelhas) de energia $10^{19}eV$ simulados pelo método de Monte Carlo. [18]	22
2.6	Mapa do detector híbrido do Observatório Pierre Auger. Os ângulo de detecção dos telescópios de fluorescência estão marcados como linhas azuis, e as posições dos detectores de superfície como pontos vermelhos. [20]	23
2.7	Representação da detecção híbrida do chuveiro atmosférico, apresentando o plano de detecção por um telescópio de fluorescência e os tanques do detector de superfície atingidos pelas partículas que chegam ao solo. [21]	23
2.8	Telescópio de fluorescência do Observatório Pierre Auger. [21]	24
2.9	Tanque do detector de superfície do Observatório Pierre Auger. [21]	25
3.1	Representação da incerteza do lançamento de um dado com distribuição de probabilidade uniforme, onde cada resultado é visto como um evento isolado, e as probabilidades de cada resultado são representados adjacentes às arestas.	28
3.2	Representação da incerteza do lançamento de um dado com distribuição de probabilidade uniforme, onde os resultados são separados nas partições A , B e C , e as probabilidades a posteriori de cada a partir de cada nó são representadas adjacentes às arestas.	29

3.3	Exemplo de superfície de uma função objetivo de duas variáveis a ser maximizada, em que o valor da função é representado em cores e no eixo vertical do gráfico.	34
3.4	Representação da aplicação do gradiente ao exemplo de função objetivo da figura anterior; a orientação das setas indica o sentido do gradiente, e seu tamanho indica a magnitude.	34
4.1	Exemplo de dados de <i>output</i> da simulação do chuveiro extenso com um próton de $10^{17}eV$ como primário gerado com o CORSIKA [3], apresentando o perfil longitudinal do número de diferentes espécies de partículas a cada $10g \cdot cm^{-2}$ de profundidade do chuveiro.	43
4.2	Comparação em escala linear e escala logarítmica entre os perfis longitudinais de partículas carregadas de duas simulações do CORSIKA [3] com as mesmas condições iniciais, sendo o primário um próton de $10^{17}eV$; a primeira simulação gerou uma colisão inicial a uma profundidade de $5 \pm 5g \cdot cm^{-2}$, enquanto na segunda ela ocorreu em $155 \pm 5g \cdot cm^{-2}$	44
4.3	Aplicação do método da entropia máxima à entropia total das distribuições por profundidade de: número de partículas eletrônicas carregadas (a), número de partículas muônicas carregadas (b), energia de ionização por partículas eletrônicas carregadas (c), e energia de ionização por partículas muônicas carregadas (d). Os pontos azuis correspondem aos dados, e as curvas vermelhas representam as distribuições obtidas.	46
4.4	Resultado da aplicação do MEM em dados com profundidade até aproximadamente $1000g \cdot cm^{-2}$ considerando a entropia da distribuição parametrizada até aproximadamente $3000g \cdot cm^{-2}$, com apenas as condições de contorno de média linear e média logarítmica. Os dados são apresentados em azul, e a distribuição obtida é representada pela curva vermelha, com pontos nas posições correspondentes às quais há a presença de dados.	48
4.5	Distribuição de Gaisser-Hillas para um valor fixo do parâmetro de forma b , diferentes valores do parâmetro de escala c , e uma profundidade primária de $30g \cdot cm^{-2}$. A figura superior tem o eixo y em escala linear, enquanto a inferior em escala logarítmica. Este eixo não apresenta valores em ambas escalas, visto que as três curvas sofreram diferentes normalizações para que seus máximos fossem iguais, meramente com o propósito de facilitar a comparação visual.	50

4.6	Distribuição de Gaisser-Hillas para diferentes valores do parâmetro de forma b , um valor fixo do parâmetro de escala c , e uma profundidade primária de $30g \cdot cm^{-2}$. A figura superior tem o eixo y em escala linear, enquanto a inferior em escala logarítmica. Este eixo não apresenta valores em ambas escalas, visto que as três curvas sofreram diferentes normalizações para que seus máximos fossem iguais, meramente com o propósito de facilitar a comparação visual.	51
4.7	Gráficos bidimensionais de estatísticas obtidas a partir da distribuição de Gaisser-Hillas em função dos parâmetros de forma b , de escala c , e uma profundidade inicial nula, comparados aos dados de um perfil longitudinal simulados com o CORSIKA [3]. São representados a média do logaritmo da profundidade (superior esquerda), o módulo do desvio relativo desta média logarítmica com relação aos dados com $\overline{\log(x)}_{data} \approx 6.5$ (inferior esquerda), a média da profundidade (superior direita), e o módulo do desvio relativo desta média com relação aos dados com $\bar{x}_{data} \approx 680$ (inferior direita). Todos os valores de profundidade em $g \cdot cm^{-2}$.	52
4.8	Distribuição de Gaisser-Hillas para diferentes valores dos parâmetros de forma b , do parâmetro de escala c , e uma profundidade primária de $30g \cdot cm^{-2}$. Os parâmetros b e c foram ajustados para que todas distribuições atendam as condições de contorno. A figura superior tem o eixo y em escala linear, enquanto a inferior em escala logarítmica. Este eixo não apresenta valores em ambas escalas, visto que as três curvas sofreram diferentes normalizações para que seus máximos fossem iguais, meramente com o propósito de facilitar a comparação visual.	53
4.9	Gráficos bidimensionais de estatísticas obtidas a partir da distribuição de Gaisser-Hillas em função dos parâmetros de forma b , de escala c , e uma profundidade inicial nula, comparados aos dados de um perfil longitudinal simulados com o CORSIKA [3]. São representados a média da profundidade (superior esquerda), e o módulo do desvio relativo desta média com relação aos dados com $\bar{x}_{data} \approx 680$ (inferior esquerda), a variância da profundidade (superior direita), e o módulo do desvio relativo desta variância com relação aos dados com $\overline{\sigma^2}_{data} \approx 28 \cdot 10^3$ (inferior direita). Todos os valores de profundidade em $g \cdot cm^{-2}$.	55

4.10	Resultado da aplicação do MEM em dados de colisão primária “rasa”, com simulação da cascata de partículas até uma profundidade de aproximadamente $1000g \cdot cm^{-2}$, considerando a entropia da distribuição parametrizada até aproximadamente $3000g \cdot cm^{-2}$, com as condições de contorno de média linear, média logarítmica, e variância dos dados. Os dados são apresentados em azul, e a distribuição obtida é representada pela curva vermelha, com pontos nas posições correspondentes às quais há a presença de dados.	56
4.11	Resultado da aplicação do MEM em dados de colisão primária comparativamente profunda, com simulação da cascata de partículas até uma profundidade de aproximadamente $1000g \cdot cm^{-2}$, considerando a entropia da distribuição parametrizada até aproximadamente $3000g \cdot cm^{-2}$, com as condições de contorno de média linear, média logarítmica, e variância dos dados. Os dados são apresentados em azul, e a distribuição obtida é representada pela curva vermelha, com pontos nas posições correspondentes às quais há a presença de dados.	57
4.12	Resultado da parametrização da profundidade primária para diversos chuveiros atmosféricos. O eixo horizontal corresponde à profundidade primária real e o eixo vertical corresponde à profundidade primária obtida pelas formas de parametrização da distribuição comparadas. Os pontos azuis representam o resultado do ajuste de curva tradicional da função de distribuição de Gaisser-Hillas realizado automaticamente pelo CORSIKA, os pontos vermelhos os resultados obtidos pela aplicação do MEM com as técnicas e condições de contorno descritas neste capítulo, e a linha verde representa a região “ideal”, em que as profundidades primárias obtidas corresponderiam perfeitamente àsquelas das dos dados simulados.	58

SUMÁRIO

Lista de Figuras	8
1 Introdução	13
2 Chuveiros atmosféricos extensos	15
2.1 Raios cósmicos	15
2.2 Características e detecção	19
2.3 Observatório Pierre Auger	22
2.4 Simulação de chuveiros atmosféricos extensos com o CORSIKA	26
3 Otimização matemática	27
3.1 Entropia da informação	27
3.2 MEM - Método da entropia máxima	30
3.3 Método dos multiplicadores de Lagrange	32
3.4 Aplicação dos métodos na parametrização da distribuição longitudinal . . .	36
3.5 Método do Lagrangiano aumentado	39
3.6 Otimização numérica em programação não-linear com o ALGENCAN . . .	40
4 Resultados e discussões	42
4.1 Dados da simulação de chuveiros com o CORSIKA	42
4.2 Aplicação do MEM para múltiplas distribuições vinculadas, e estudo dos perfis eletrônico e muônico de número e energia	45
4.3 Aplicação do MEM fora da região dos dados conhecidos	46
4.4 Estudo de máximos locais da entropia a partir dos parâmetros de forma e escala	49
4.5 Introdução da terceira condição de contorno	55
4.6 Determinação da profundidade da colisão primária	58
5 Conclusões e perspectivas	60
Referências	61

1. INTRODUÇÃO

Na área de física de astropartículas alguns dos experimentos de destaque visam detectar raios cósmicos de energia acima de $10^{18}eV$ [4, 5]. A detecção desses raios cósmicos se dá através da observação de chuveis atmosféricos extensos – tipicamente é criada uma cascata de até 10^{11} partículas secundárias espalhadas por áreas de até $20km^2$.

No observatório Pierre Auger [6], esses chuveis são observados com telescópios desenvolvidos para detecção da fluorescência e da radiação Cherenkov emitidos pela interação das partículas com o nitrogênio da atmosfera e com tanques de detecção de Cherenkov que detectam diretamente a passagem das partículas que chegam até a superfície da Terra. Da observação do chuveiro através da atmosfera é extraído o perfil longitudinal, enquanto da detecção na superfície obtém-se o perfil lateral.

Foco deste trabalho, a parametrização do perfil longitudinal é uma das ferramentas cruciais para que se consiga determinar características fundamentais da partícula primária, como o ponto da colisão inicial, sua energia e sua composição. O método de parametrização do perfil irá afetar não só a sensibilidade a ruídos experimentais dos resultados, como também o grau de viés na caracterização do chuveiro devido às premissas assumidas.

Na extração de informações do perfil longitudinal é inevitável adicionar suposições aos dados medidos inerentemente incompletos. O perfil é limitado pela distância disponível entre a origem do chuveiro atmosférico e o solo, o ruído da mensuração e a incerteza estatística; a própria escolha de uma função de distribuição para o ajuste do perfil é uma informação externa aos dados mensurados. É inevitável inserir premissas junto aos dados obtidos na extração de qualquer informação na qual haja incerteza.

Nesse contexto, o grau de suposições inseridas junto aos dados de um problema pode ser medido com o conceito de *entropia da informação*, explorado pela *teoria da informação* [7], e relacionado intimamente com os conceitos de entropia da termodinâmica e da mecânica estatística. Em suma, quanto maior o valor de entropia da parametrização de uma distribuição, menor será a interferência das suposições inseridas junto aos dados.

Este trabalho propõe o uso do *método da entropia máxima* (MEM) [1, 2] como uma alternativa de parametrização do perfil longitudinal para minimizar o enviesamento dos dados devido à própria escolha de medidas utilizadas e da forma da distribuição. Assim o método reduz o viés intrínseco das premissas assumidas no ajuste de curvas de distribuição aos dados.

No capítulo 2 é apresentado o panorama do estudo chuviros atmosféricos extensos. São abordadas as características dos raios cósmicos, o mistério dos raios cósmicos de ultra-altas energias, o observatório Pierre Auger [6], e os métodos empregados na detecção de chuviros atmosféricos e em seu estudo. Entre essas é destacada a análise dos perfis longitudinal e lateral, e a simulação da interação das partículas com a atmosfera pelo método de Monte Carlo utilizando o programa CORSIKA [3].

O capítulo 3 descreve toda a metodologia de otimização matemática utilizada neste trabalho. Ele parte do desenvolvimento do conceito de *entropia da informação* e da motivação do *método da entropia máxima* [1, 2], utilizando então o *método dos multiplicadores de Lagrange* [8] para poder definir a forma da função de distribuição a partir dos dados experimentais de interesse. Ele também descreve o *método do Lagrangiano aumentado* utilizado pelo programa de otimização ALGENCAN [9], que é utilizado para realizar o tratamento numérico utilizado para a obtenção de resultados deste trabalho.

O capítulo 4 expõe os resultados da utilização da otimização matemática, descrita no capítulo anterior, na parametrização do perfil longitudinal de chuviros atmosféricos simulados com o CORSIKA [3]. São apresentados os resultados referentes à componente eletrônica, à componente muônica, e uma análise de como os parâmetros obtidas para essas duas componentes se relacionam.

O capítulo 5 apresenta as conclusões do trabalho e propõe outros uso dos métodos explorados no estudo de chuviros atmosféricos extensos.

2. CHUVEIROS ATMOSFÉRICOS EXTENSOS

Por quase um século, o entendimento de raios cósmicos de baixa e médias energias progrediram muito, porém os raios cósmicos de ultra-altas energias (superiores a $10^{18}eV$) são objetos de muitos estudos, nos quais estão envolvidos os processos mais energéticos do universo. A solução dos mistérios envolvendo essas partículas pode aumentar a compreensão sobre os possíveis corpos celestes envolvidos em sua produção e propagação até a Terra, e até mesmo ampliar os horizontes da física de partículas para além do modelo padrão. Além disso, raios cósmicos ultra-energéticos sofrem relativamente pequenas deflexões. São bem conhecidos os campos magnéticos da Via Láctea e as trajetórias dos raios cósmicos podem ser reconstruídas; como fora da galáxia os campos magnéticos são menos conhecidos, a reconstrução de trajetórias é menos precisa. A partir das reconstruções, é possível concluir que raios cósmicos devem ter origem próxima; as reconstruções mais precisas de trajetórias, realizadas com os dados do observatório Pierre Auger, indicam origens a distâncias de até 100 Mpc, enquanto outros raios cósmicos para os quais não foi possível realizar a reconstrução com a mesma precisão indicam distâncias de até 200 Mpc.

2.1 Raios cósmicos

Os raios cósmicos que serão tratados aqui são fundamentalmente prótons e outros núcleos atômicos, além de outras partículas massivas em menor escala, que atingem a Terra com altas energias produzindo chuviros atmosféricos. O estudo geral dos raios cósmicos é feito pela observação de características como a sua energia, composição, e direção de origem que permitem investigar os mecanismos de produção e propagação dessas partículas. Seu espectro vai de aproximadamente 10^6eV até $10^{20}eV$ – para partículas com energia de até aproximadamente $10^{12}eV$, o fluxo que atinge a Terra é suficiente (cerca de 1 partícula por m^2 por segundo) para que elas sejam detectadas diretamente em altas altitudes com uso de satélites ou balões. Por outro lado a medição direta de partículas com energias superiores a $10^{15}eV$ é muito difícil, dado os fluxos que podem ser inferiores a 1 partícula por m^2 por ano. Porém raios cósmicos com energias superiores a $10^{14}eV$ produzem reações nucleares na atmosfera que dão origem a cascatas de partículas secundárias. Essa cascata atravessa toda a atmosfera até atingir a superfície do planeta, conhecida como chuva

atmosférico extenso. A sua mensuração com o uso de aparatos de detecção localizados na superfície podem ser utilizados como uma forma de observação indireta da partícula inicial, denominada partícula primária ou simplesmente *primário*.

Uma das formas mais usadas para a caracterização dos raios cósmicos é o seu espectro de energia, apresentado na figura 2.1. Ele pode ser descrito pela lei de potência

$$F(E) = A \cdot E^{-a} \quad (2.1)$$

em que F é o fluxo de partículas com energia E . É notável que a lei de potência seja capaz de descrever aproximadamente dados experimentais de partículas cujos fluxos são tão alto quanto 10^4 partículas por $m^2 \cdot sr \cdot s$ e tão baixos quanto 10^{-28} partículas por $m^2 \cdot sr \cdot s$.

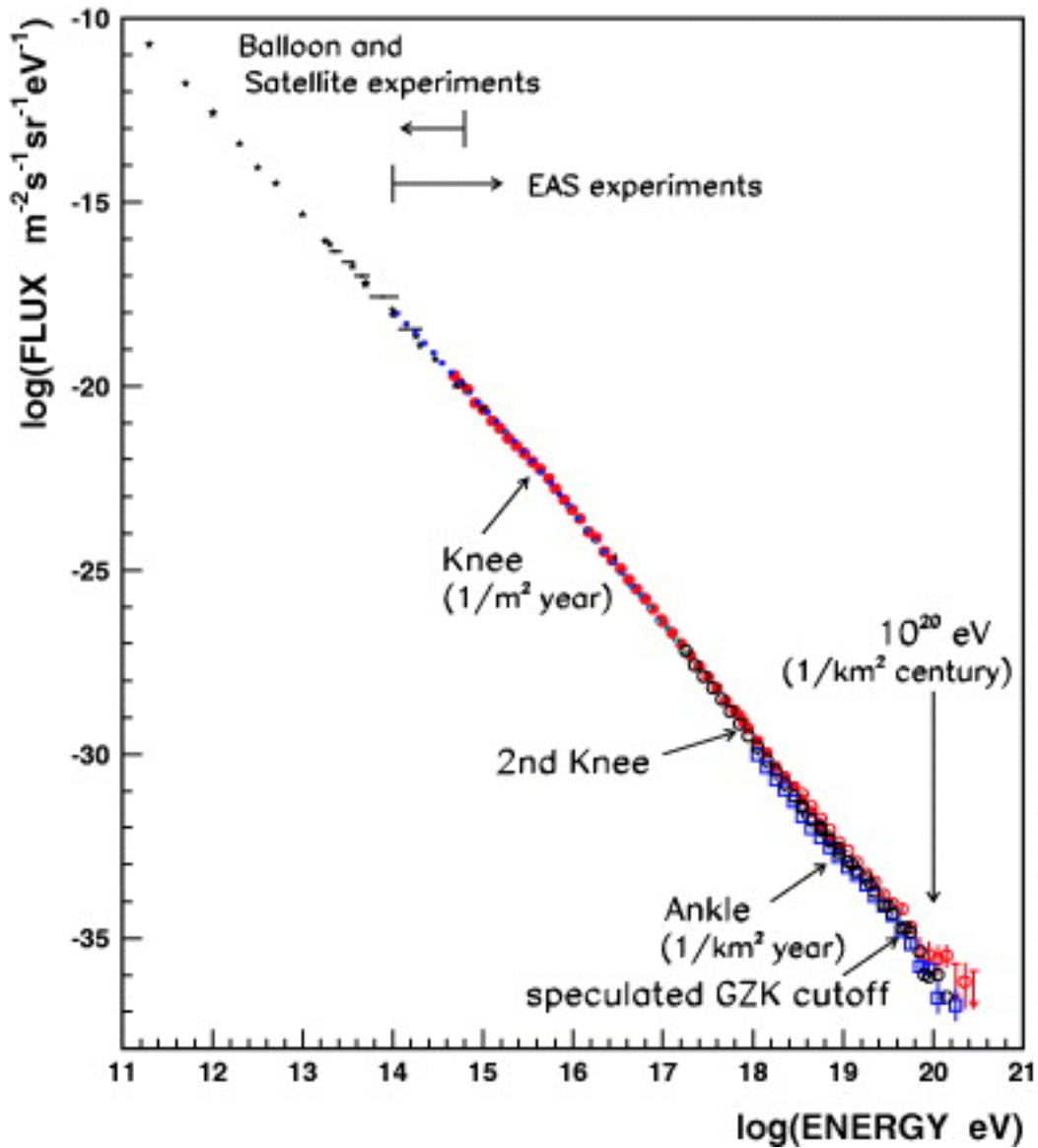


Figura 2.1: Espectro normalizado de energia de raios cósmicos. [10]

A parametrização deste espectro tem um resultado que é aproximadamente

$$F(E) = 2.7 \cdot E^{-2.7} \quad (2.2)$$

entretanto existem duas regiões que divergem da lei de potência simples, e que são particularmente relevantes para a física e a astrofísica: dois “joelhos” e um “tornozelo”, apontados na figura 2.1.

O primeiro joelho aparece na região de energia em torno de $3 \cdot 10^{15} \text{ eV}$ (a partir de onde o espectro cai mais rapidamente em função da energia, com dependência $E^{-3.1}$), o segundo joelho encontra-se próximo de $5 \cdot 10^{17} \text{ eV}$ (com $E^{-3.3}$), e o tornozelo ocorre perto de $3 \cdot 10^{18} \text{ eV}$ (onde a inclinação espectral sobe novamente a um valor próximo àquele da primeira região do espectro). [11]

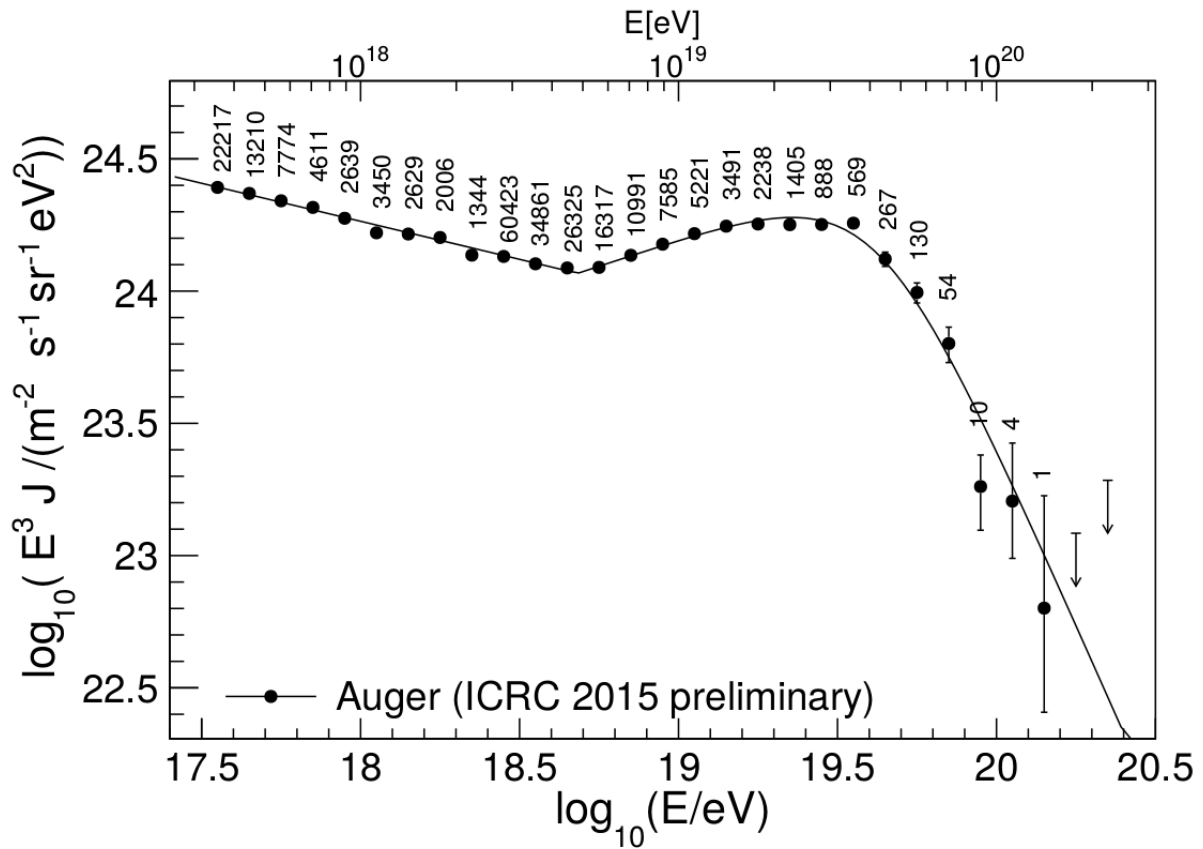


Figura 2.2: Espectro normalizado de energia de raios cósmicos na faixa de ultra-altas energias [12]

É na região de menor energia que se encontra a maior quantidade de dados disponíveis para serem analisados, visto o alto fluxo de 1 partícula por $\text{m}^2 \cdot \text{s}$, que permitem o estudo direto da composição química do primário. Para essa faixa de energia que vai até quase o primeiro “joelho”, possui-se então uma boa base de conhecimento sobre os raios cósmicos e suas fontes intra-galáticas (origem de uma fração significativa dos mesmos).

Há indícios que os raios cósmicos originados na Via Láctea se propagam por difusão e aqueles com energias mais altas tenham sido acelerados nas camadas externas das ondas de choque de remanescentes de supernovas. Os raios cósmicos carregados sofrem influência de campos magnéticos gerados por esses eventos, sendo maiores os desvios de aceleração quanto menor a energia da partícula. Assim as trajetórias de primários de energias mais baixas são de difícil reconstrução, tornando mais improvável a determinação de seu ponto de origem.

As partículas com energias superiores a $10^{18}eV$ são denominadas *raios cósmicos de ultra-altas energias*, comumente chamados na literatura pela sigla UHECR (*ultra-high-energy cosmic rays*). A interferência exercida por campos magnéticos com esse nível de energia é pouco significativa, mesmo para os núcleos mais pesados (com maiores cargas), tal que eles seriam capazes de escapar de suas galáxias de origem e percorrer grandes distâncias intergaláticas. Adicionalmente, não são conhecidas fontes astrofísicas de interação dentro da Via Láctea capazes de acelerar partículas até essas energias, tal que uma das hipóteses acerca da fonte dos raios cósmicos de ultra-alta energia é que eles tenham origem extragaláctica.

Mas essa hipótese também tem problemas – ao longo da propagação das partículas até a atmosfera da Terra, elas sofrem uma série de interações não só com outras partículas e campos magnéticos, mas também com a *radiação cósmica de fundo em microondas*. Essa radiação foi descoberta experimentalmente em 1965 por Penzias e Wilson [13], e em 1966 Greisen [14], Zatsepin e Kuz'min [15] determinaram um limite superior teórico para a energia com qual raios cósmicos com origem distante poderiam chegar ao planeta devido sua interação com o fundo de microondas. Chamado de corte GZK em homenagem aos seus proponentes, ele limitaria em particular prótons com origem em outras galáxias a energias inferiores a $5 \cdot 10^{19}eV$, vista a produção de pares de píons na interação entre os núcleos e os fótons de microondas que reduziria a energia de raios cósmicos acima do limite.

Dessa forma, os raios cósmicos observados com energias que chegam acima de $10^{20}eV$ também não poderiam ter origem em outras galáxias; por outro lado, não há conhecimento de objetos próximos à Via Láctea que possam produzir ou acelerar partículas até energias tão elevadas.

Assim existem várias perguntas ainda não respondidas no estudo dos UHECR: a partir de qual região do espectro de energia está localizada a sua componente extragaláctica? Como raios cósmicos atingem tais energias? Quais objetos astrofísicos são capazes de acelerá-los? Qual o limite de energia do espectro?

Isso leva a crer que seja possível que alguns dos fenômenos astrofísicos mais interessantes ainda sejam desconhecidos. A origem dos UHECR poderia revelar novos objetos astronômicos, ou até mesmo propriedades ainda desconhecidas da física de partículas.

2.2 Características e detecção

O chuvaireo atmosférico gerado por um raio cósmico incidente na Terra pode ser representado como um disco de partículas relativísticas que percorre a atmosfera. Pode-se observar o seu desenvolvimento na atmosfera a partir da luz de fluorescência atmosférica gerada pela interação das partículas do chuvaireo com moléculas de nitrogênio e da luz gerada pelo efeito Cherenkov daquelas com velocidades mais elevadas; também é possível detectar as partículas que chegam à superfície da Terra conforme elas atravessam detectores localizados no solo. Há duas formas muito utilizadas de avaliar o chuvaireo atmosférico, seu perfil longitudinal e seu perfil lateral; o perfil longitudinal é formado pela medida do número de partículas ou da energia depositada em função da profundidade de propagação ao longo da atmosfera, e o perfil lateral é formado por medidas da mesma natureza, mas restritas à superfície do planeta, em função da distância ao eixo longitudinal do chuvaireo.

Muitos avanços foram realizados na tecnologia de detecção de chuvaireos atmosféricos – os primeiros detectores utilizados foram câmaras de ionização, contadores Geiger-Müller e câmaras de nuvens, que somente permitem detecção na passagem de partículas carregadas pelo seu interior. Assim eles possibilitam observar as partículas apenas quando elas chegam ao solo, ou em pontos isolados com o auxílio de balões e satélites, tal que sem outros recursos seja impossível avaliar o desenvolvimento da cascata conforme ela atravessa a atmosfera.

O desenvolvimento de telescópios sensíveis à luz emitida por fluorescência e pelo efeito Cherenkov pela interação das partículas com o gás atmosférico gerou contribuições extremamente relevantes para a observação dos chuvaireos. As partículas carregadas ionizam moléculas do gás durante a sua passagem, e as moléculas de nitrogênio em particular sofrem um processo de excitação de elétrons em níveis moleculares que deexcitam na forma de fluorescência em comprimentos de onda na faixa do ultravioleta próximo. Aquelas partículas carregadas que se propagam pelo gás com velocidades superiores à da luz no meio permitem sua detecção em ângulos bem determinados devido ao efeito Cherenkov. Entretanto a luz emitida por estes processos não é particularmente intensa, tal que sua

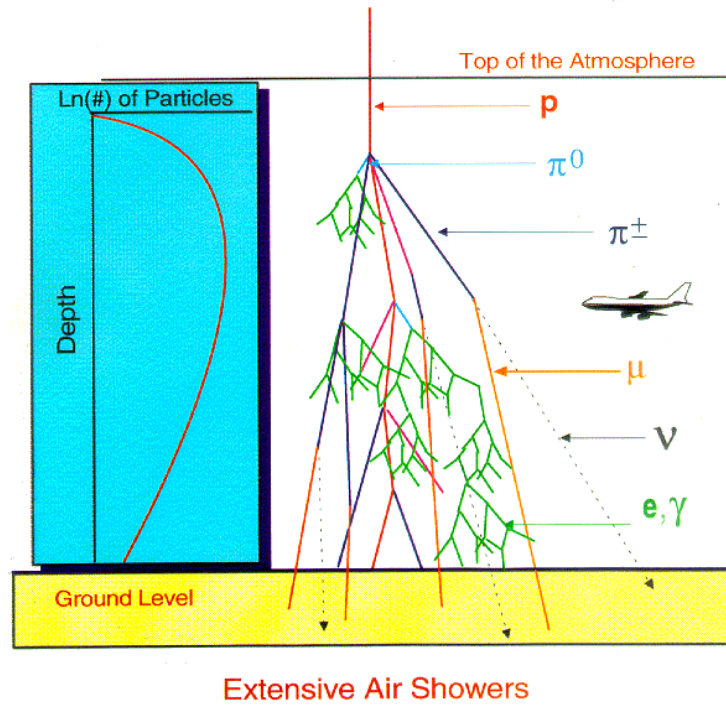


Figura 2.3: Representação do desenvolvimento de um chuveiro atmosférico, com o logaritmo do número de partículas em função da profundidade na atmosfera à esquerda e as interações típicas do primário e dos secundários à direita. [16]

deteção é limitada por condições como a posição e fase da Lua, a umidade relativa do ar e a presença de nuvens.

A posição do máximo de partículas no perfil longitudinal do chuveiro atmosférico é um dos parâmetros importantes na determinação da composição química do primário, enquanto a análise da energia total do chuveiro fornece a sua energia inicial. A análise dos eventos que formam o desenvolvimento da cascata de partículas permitem que sejam avaliados modelos de interação entre partículas em níveis de energia muito além daqueles obtíveis em colisores de partículas, que atualmente estão na faixa de 10^{16}eV .

Para identificar chuveiros atmosféricos extensos de diferentes primários, as grandezas mensuradas experimentalmente precisam ser comparados com resultados de simulações pelo método de Monte Carlo. Essas simulações tratam o transporte e as interações das partículas com a matéria, com base no conhecimento atual de física de partículas e sustentados por experimentos realizados em laboratórios com condições controladas. A interpretação da medida de um chuveiro atmosférico extenso depende, portanto, dos modelos teóricos utilizados na simulação. Enquanto as interações eletromagnéticas (responsáveis pela componente eletromagnética e as interações de ionização, produção de fluorescência, efeito Cherenkov, etc) e interações fracas (decaimentos de partículas instáveis) são bem

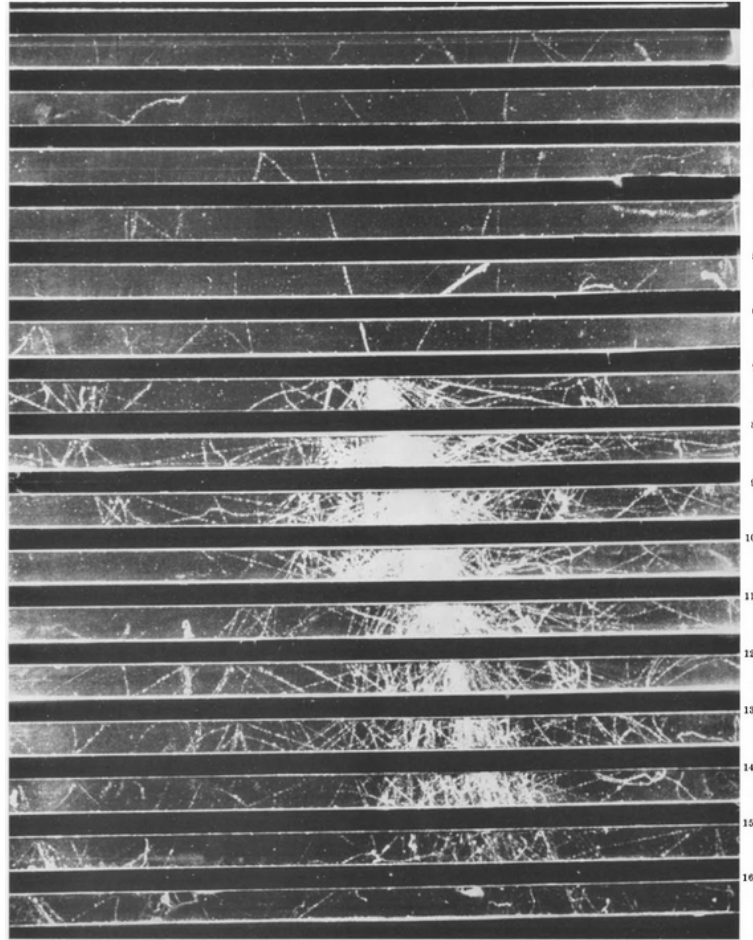


Figura 2.4: Desenvolvimento de cascata a 3027m de altitude em uma câmara de nuvens que contém placas de chumbo de 1.3cm de espessura (faixa escuras numeradas). A área da seção de choque da câmara era de $0.5 \times 0.3\text{m}^2$. [17]

estabelecidas pela QED (*quantum electrodynamics*), as interações hadrônicas ainda são as maiores fontes de incertezas desses modelos.

A maior parte dos modelos de interação hadrônica utilizados nas simulações de chuueiros foram construídos a partir de resultados experimentais. Estes experimentos foram realizados predominantemente criando colisões próton-próton em energias que atingem até a ordem de 10^{15} e 10^{16}eV em laboratório; os dados resultantes são então extrapolados para prever as interações em faixas de energia superiores.

Outra fonte de informações sobre a composição química do primário e das interações que ocorrem durante o desenvolvimento do chuveiro é a relação entre as componentes eletrônica e muônica das partículas secundárias. Entretanto, os resultados teóricos dos diferentes modelos utilizados nas simulações apresentam divergências que causam um impacto significativo na conexão entre a relação dessas componentes e a composição química do primário, em particular para núcleos pesados. Uma série de estudos vem

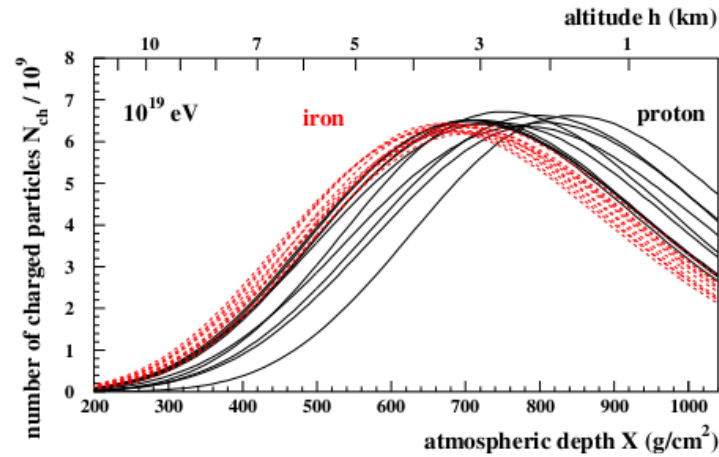


Figura 2.5: Perfis longitudinais de número de partículas carregadas para primários de próton (em linhas sólidas pretas) e ferro (em linhas tracejadas vermelhas) de energia $10^{19} eV$ simulados pelo método de Monte Carlo. [18]

sendo realizada para entender como as diferenças entre os modelos criam divergências no resultado de simulações de chuueiros atmosféricos extensos [19].

2.3 Observatório Pierre Auger

Localizado na Argentina próximo à cidade de Malargüe, na planície *Pampa Amarilla*, o Observatório Pierre Auger ocupa cerca de $3000 km^2$ com um “detector híbrido” capaz de observar chuueiros atmosféricos extensos ao longo de sua passagem na atmosfera e ao atingir o solo.

O observatório, cujo nome é uma homenagem ao físico pioneiro no estudo de chuueiros atmosféricos extensos Pierre Victor Auger, tornou-se em 2003 o maior detector de raios cósmicos de energias ultra-altas do mundo. Sua forma de detecção híbrida é composta por 27 telescópios de fluorescência em quatro posições do observatório e 1660 detectores de superfície distribuídos na área do experimento.

A luz ultravioleta emitida devido à interação das partículas da cascata com o nitrogênio da atmosfera forma um rastro que pode ser detectado por telescópios que fazem parte do detector de fluorescência – a propagação do chuueiro atmosférica pode ser imaginada como um clarão que começa a ser emitido da região da colisão primária e vai se tornando mais brilhante conforme se propaga pelo volume da cascata à velocidade da luz. Entretanto, a luz emitida pelo chuueiro tem energia inferior a 4 watts, tal que mesmo que seres humanos pudessem enxergar na faixa do ultravioleta o brilho seria muito pequeno para

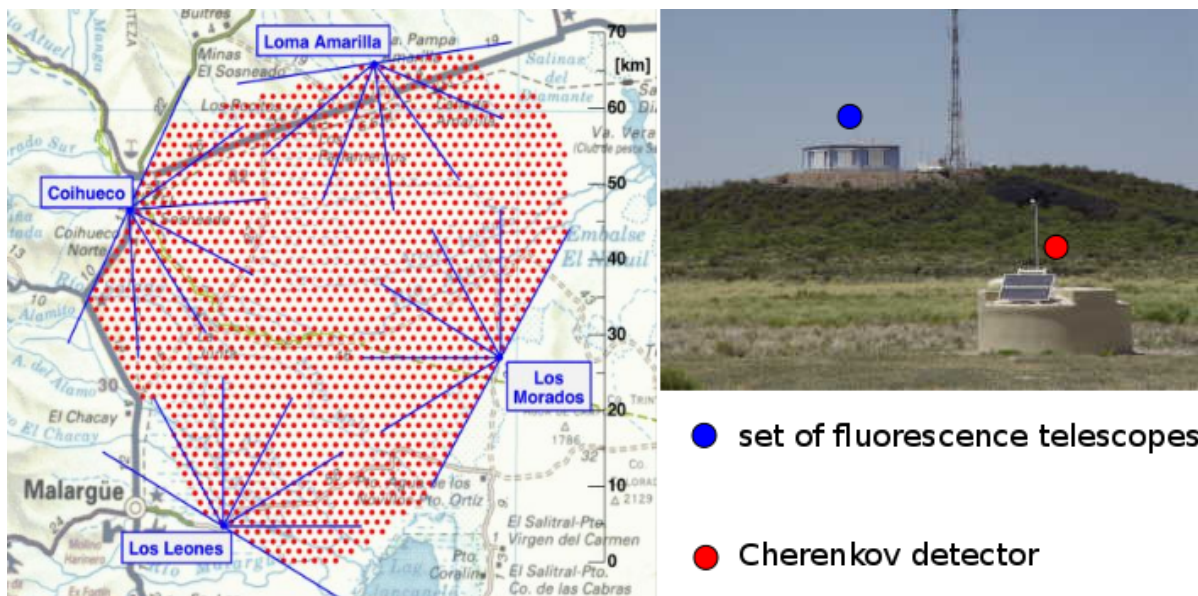


Figura 2.6: Mapa do detector híbrido do Observatório Pierre Auger. Os ângulo de detecção dos telescópios de fluorescência estão marcados como linhas azuis, e as posições dos detectores de superfície como pontos vermelhos. [20]

ser perceptível; mesmo para os telescópios que tem sensibilidade muito maior à do olho humano, é indispensável que as medições sejam realizadas distantes de fontes de poluição luminosa, como cidades. Até mesmo o brilho da Lua é suficiente para interferir com as medidas, assim como a presença de umidade no ar, tal que as medições de fluorescência precisam ser realizadas sempre em função das condições astronômicas e ambientais.

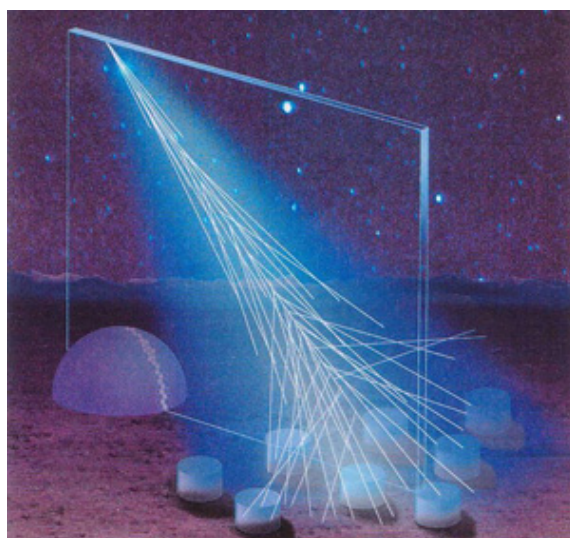


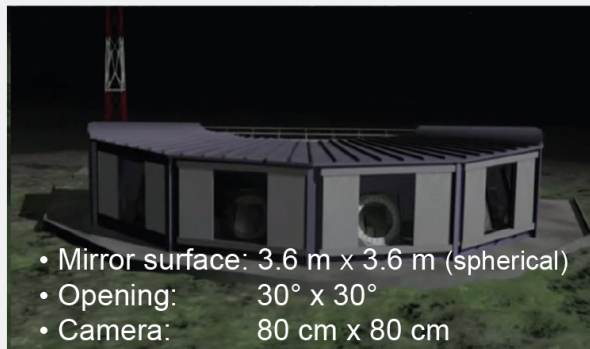
Figura 2.7: Representação da detecção híbrida do chuveiro atmosférico, apresentando o plano de detecção por um telescópio de fluorescência e os tanques do detector de superfície atingidos pelas partículas que chegam ao solo. [21]

Os telescópios utilizam um *grid* de espelhos côncavos para coletar luz suficiente em sensores, em um aparato capaz de “enxergar” chuveiros atmosféricos a até 15 km de

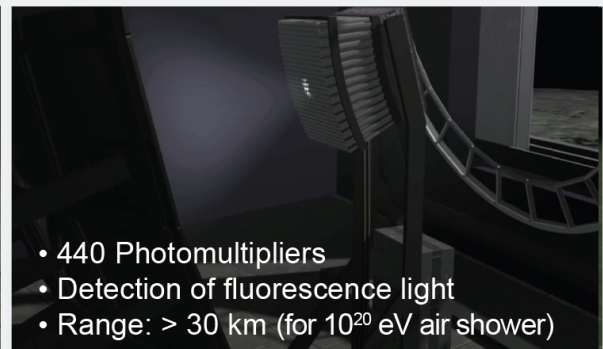
distância. O conjunto de telescópios de fluorescência é o aparato experimental utilizado para obter os dados do perfil longitudinal dos chuveiros, que é o objeto de interesse desta dissertação.

Fluorescence Detector

27 fluorescence telescopes
(in 4 different places)



- Mirror surface: 3.6 m x 3.6 m (spherical)
- Opening: 30° x 30°
- Camera: 80 cm x 80 cm



- 440 Photomultipliers
- Detection of fluorescence light
- Range: > 30 km (for 10^{20} eV air shower)



Figura 2.8: Telescópio de fluorescência do Observatório Pierre Auger. [21]

O detector de superfície é composto por 1660 tanques de água de 3.6 metros de diâmetro distribuídos na área de cerca de 3000 km² do experimento. Cada tanque é separado de seus vizinhos por uma distância de 1.5 km e contém 12 mil litros de água. A luz Cherenkov produzida por partículas que atravessam a água em velocidade superior à da luz neste meio é detectada por fotomultiplicadoras no interior do tanque, que medem a quantidade de luz

emitida. A comparação dessa intensidade e os instantes de detecção entre os múltiplos tanques permitem detectar não só o ponto onde o eixo longitudinal do chuveiro atmosférico intercepta o solo, mas também a superfície de propagação da frente do chuveiro. Essas informações permitem a descrição completa da posição do eixo longitudinal de propagação no volume atmosférico.

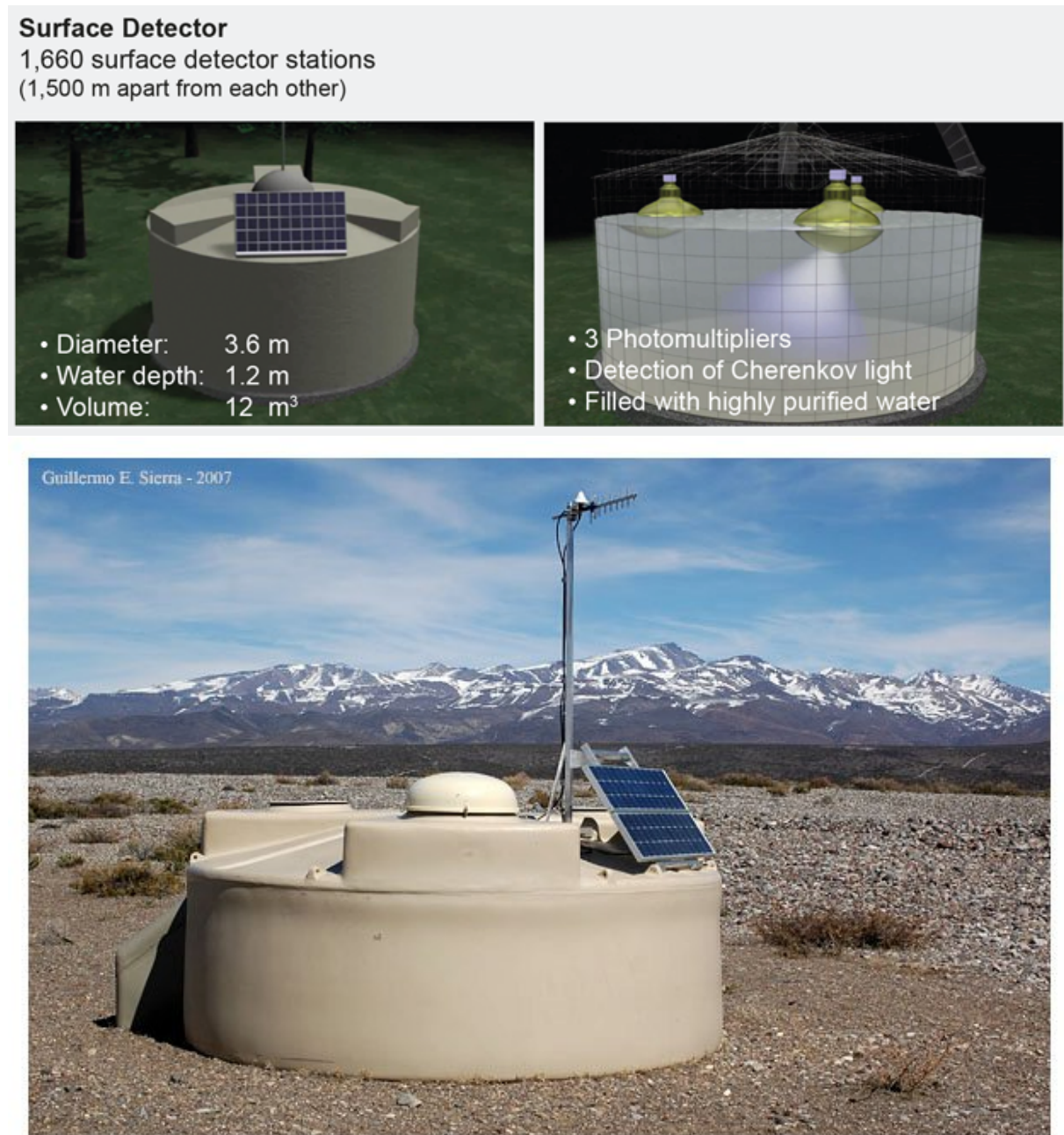


Figura 2.9: Tanque do detector de superfície do Observatório Pierre Auger. [21]

O conjunto de tanques do detector de superfície é o aparato experimental também utilizado para obter os dados do perfil lateral dos chuveiros.

2.4 Simulação de chuueiros atmosféricos extensos com o CORSIKA

As simulações dos chuueiros atmosféricos extensos foram realizadas utilizando-se o CORSIKA (*Cosmic Ray Simulations for Kaskade*) [3], um programa de simulação baseado no método de Monte Carlo, o mais usado pela comunidade de física de raios cósmicos. Com este programa é possível simular chuueiros atmosféricos extensos cujos primários são prótons, núcleos pesados, fótons, entre outras partículas. O intervalo de energia dos primários simulados vai de $10^{12}eV$ até $10^{20}eV$.

Todos os processos físicos relevantes para o desenvolvimento da cascata de partículas na atmosfera são simulados, envolvendo todas as interações fortes, fracas e eletromagnéticas previstas nos modelos de cascata na atmosfera. O primário e as partículas secundárias têm suas trajetórias individuais desenvolvidas ao longo da atmosfera simulada, sendo seus parâmetros alterados conforme interações regidas pelas seções de choque relativas às características das partículas e da região atmosférica, até chegarem ao nível do solo. Neste ponto suas informações são armazenadas para que as partículas sejam tratadas com relação à detecção na superfície. Desta maneira espera-se que a simulação seja tão fiel quanto os modelos permitirem para descrever as características e flutuações estatísticas de eventos reais.

As interações eletromagnéticas de um chuueiro atmosférico são tratadas pela rotina EGS4 (*Electron Gamma Shower System version 4*) [22]. Uma parte crítica, devido à inevitável extrapolação de dados para partículas de altas energias, é a simulação das interações hadrônicas; no CORSIKA existem diferentes opções de modelos disponíveis para esse tipo de interação, como os conhecidos QGSJet (*Quark Gluon String Model with Jets*) [23], SIBYLL [24], e EPOS LHC [25].

3. OTIMIZAÇÃO MATEMÁTICA

3.1 Entropia da informação

Em um artigo publicado em 1948 [7], Claude Shannon realizou uma análise da estrutura estatística da informação contida em mensagens. Este trabalho inaugurou todo um novo campo de estudo chamado de *teoria da informação*, que vem se tornando extremamente relevante em uma era na qual a quantidade de informação disponível cresce cada vez mais rápido, atingindo as ciências naturais, as ciências formais e as ciências sociais.

O trabalho revolucionário de Shannon apresentou o conceito de *entropia da informação* como uma medida do grau de incerteza de uma fonte de informação. Em uma função de distribuição de probabilidade discreta uniforme, por exemplo, a incerteza cresce conforme aumenta-se o número de estados possíveis. Por outro lado, fixando-se o número de estados de uma distribuição não uniforme, quanto mais “desbalanceada” for a distribuição, aumentando a probabilidade de se obter um conjunto pequeno de estados e diminuindo a probabilidade dos outros resultados, menor será a incerteza.

Em seu artigo [7], Shannon usa como exemplo uma mensagem escrita na língua inglesa: em qualquer texto escrito, as letras se organizarão formando fonemas válidos da língua, partes de um vocabulário finito de palavras. Logo qualquer mensagem ortograficamente correta possui um grau de incerteza inferior a uma distribuição aleatória de letras e espaços. O próximo passo então é definir uma forma de mensuração desta grandeza, que não só seja relevante em significado mas que também seja internamente consistente. Procura-se então uma medida do grau de incerteza S de um sistema discreto que seja função das probabilidades p_i dos n estados possíveis, na forma $S(p_1, p_2, \dots, p_n)$, que possua as seguintes propriedades:

1. S é uma função contínua de p_i .
2. Se todos os p_i são iguais (distribuição uniforme), então S deve ser uma função monotônica crescente em função de n , pois quanto maior o número de estados igualmente prováveis, maior será a incerteza.
3. Separando os estados em partições (A, B, \dots, Ω) , sendo a probabilidade de cada partição a soma das probabilidades dos estados contidos ($p_A = p_{A1} + p_{A2} + \dots$), o

valor original de S deve ser igual à soma dos valores de S de cada grupo isoladamente ponderados pela probabilidade do grupo, inclusive o grupo de partições:

$$\begin{aligned}
 S(p_1, p_2, \dots, p_n) &= 1 \cdot S(p_A, p_B, \dots, p_Z) \\
 &+ p_A \cdot S\left(\frac{p_{A1}}{p_A}, \frac{p_{A2}}{p_A}, \dots\right) \\
 &+ p_B \cdot S\left(\frac{p_{B1}}{p_B}, \frac{p_{B2}}{p_B}, \dots\right) \\
 &+ \dots \\
 &+ p_\Omega \cdot S\left(\frac{p_{\Omega 1}}{p_\Omega}, \frac{p_{\Omega 2}}{p_\Omega}, \dots\right)
 \end{aligned} \tag{3.1}$$

onde é óbvio que $\frac{p_{A1}}{p_A}$ é a probabilidade condicional $P(A1|A)$ de que ocorra o estado $A1$ visto que se conhece a priori que o resultado pertence à partição A .

As duas primeiras propriedades são suficientemente simples, mas a terceira pode ser ilustrada com um exemplo: um único lançamento de um dado de seis faces com distribuição de probabilidade uniforme ($p_i = \frac{1}{6}$) é um processo em que a incerteza reside no desconhecimento do resultado, que pode ser representado como na figura 3.1.

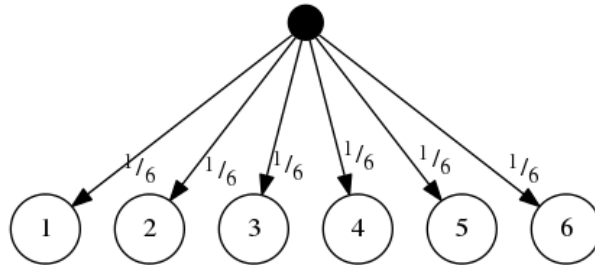


Figura 3.1: Representação da incerteza do lançamento de um dado com distribuição de probabilidade uniforme, onde cada resultado é visto como um evento isolado, e as probabilidades de cada resultado são representados adjacentes às arestas.

A incerteza desta distribuição uniforme com 6 estados pode ser representada por:

$$S\left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right) \tag{3.2}$$

Por outro lado, é admissível particionar os possíveis resultados do lançamento em grupos arbitrários; este particionamento pode ser como um processo de decisão em duas etapas, ou meramente como uma única etapa em que os resultados pertencem a classes diferentes. Neste exemplo os resultados do lançamento são agrupados como visto na figura 3.2, e as probabilidades de o resultado do lançamento pertencer a uma das partições podem ser vistas nas equações 3.3.

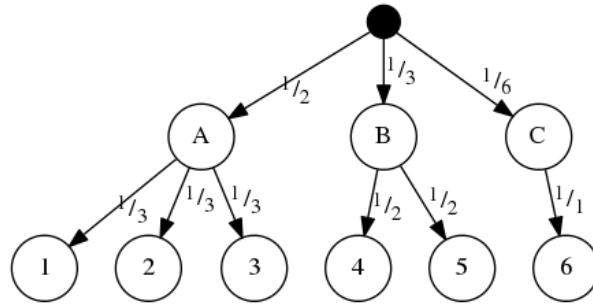


Figura 3.2: Representação da incerteza do lançamento de um dado com distribuição de probabilidade uniforme, onde os resultados são separados nas partições A , B e C , e as probabilidades a posteriori de cada a partir de cada nó são representadas adjacentes às arestas.

$$\begin{aligned} p_A &= p_1 + p_2 + p_3 = \frac{1}{2} \\ p_B &= p_4 + p_5 = \frac{1}{3} \\ p_C &= p_6 = \frac{1}{6} \end{aligned} \quad (3.3)$$

Suponha-se que seja ignorada a face resultante do lançamento do dado, observando-se somente a qual partição o resultado pertence, e a probabilidade de cada partição. Neste caso, a incerteza resultante será

$$S(p_A, p_B, p_C) = S\left(\frac{1}{2}, \frac{1}{3}, \frac{1}{6}\right) \quad (3.4)$$

Por outro lado, caso observe-se os resultados possíveis para cada partição isoladamente, ignorando-se as outras partições, então a incerteza de cada partição pode ser expressa nas formas a seguir:

$$\begin{aligned} S(p_{A1}, p_{A2}, p_{A3}) &= S\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) \\ S(p_{B1}, p_{B2}) &= S\left(\frac{1}{2}, \frac{1}{2}\right) \\ S(p_{C1}) &= S\left(\frac{1}{1}\right) \end{aligned} \quad (3.5)$$

Como foram ignoradas as outras partições, para a partição A existem apenas 3 resultados igualmente prováveis, 2 para B , e 1 para C .

Para avaliar a incerteza total do lançamento do dado considerando as três partições, deve-se então somar a incerteza do resultado da partição (equação 3.4) ponderada pela probabilidade $\frac{1}{1}$ (a incerteza do resultado de partição ocorrerá em 100% dos lançamentos) e as incertezas de cada uma das três partições (equações 3.5) ponderadas pelas probabilidades p_A , p_B e p_C (as incertezas de cada partição ocorrem em apenas respectivamente $\frac{1}{2}$, $\frac{1}{3}$ e $\frac{1}{6}$ dos lançamentos). Ou seja, considerando as equações acima, as terceira propriedade

implica que ocorra a igualdade

$$\begin{aligned}
 S\left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right) &= \frac{1}{1} \cdot S\left(\frac{1}{2}, \frac{1}{3}, \frac{1}{6}\right) \\
 &+ \frac{1}{2} \cdot S\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) \\
 &+ \frac{1}{3} \cdot S\left(\frac{1}{2}, \frac{1}{2}\right) \\
 &+ \frac{1}{6} \cdot S\left(\frac{1}{1}\right)
 \end{aligned} \tag{3.6}$$

É possível demonstrar [7] que a única função que atende as três propriedades é

$$S = -k \cdot \sum_{i=1}^n \{p_i \cdot \log(p_i)\} \tag{3.7}$$

onde k é uma constante positiva.

Não é difícil perceber que a entropia da informação tem uma forma extremamente semelhante à definição termodinâmica de entropia, em particular a *entropia de Gibbs*, e ecoa o trabalho desenvolvida em mecânica estatística por Boltzmann no desenvolvimento da *equação de Boltzmann*. Um estudo abordando a relação entre as diferentes definições de entropia foi publicado por E. T. Jaynes [26].

A constante k poderia também ser definida como a constante de Boltzmann, para ser compatível com às grandezas termodinâmicas da entropia física; entretanto como aqui o foco de estudo é a estrutura estatística da informação, e a entropia absoluta não é relevante, por simplicidade é comum definir $k := 1$ tal que a expressão de entropia se torna

$$S = - \sum_{i=1}^n \{p_i \cdot \log(p_i)\} \tag{3.8}$$

Como a troca de base do logaritmo também equivale a uma mudança do valor de k , a escolha da base também pode ser qualquer. É comum que se escolha a base neperiana na física, a base 10 em diversas engenharias, e a base 2 em áreas relacionadas com a computação.

3.2 MEM - Método da entropia máxima

Em todas as áreas de estudo em que se faça uma análise quantitativa de processos estocásticos, lida-se com a coleção de dados mensuráveis e suposições que interferirão nas medidas ou na sua análise. A amostragem de dados não somente pode ser limitada por dados incompletos (devido a limites na sua mensuração), como sempre sofrerá de erro

estatístico. Para os propósitos deste trabalho, qualquer divergência deste tipo entre os dados e a distribuição “ideal” do objeto de estudo será chamada de ruído.

Como o ruído é intrinsecamente não nulo, os dados serão sempre insuficientes para determinar univocamente a distribuição de probabilidades por trás do processo. A metodologia de elaborar uma hipótese de que os dados seguem uma função de distribuição e o ajuste desta função aos dados é externa aos dados em si, é uma suposição adicionada às evidências experimentais. A escolha de uma função para ajuste não é uma inferência dedutiva, mas sim indutiva, tal que ela configura um viés da análise do processo.

É claro que, particularmente nas ciências naturais, é impossível evitar as suposições, realizar uma análise puramente dedutiva, livre de qualquer tipo de viés, visto o próprio limite de acesso à informação intrínseco à natureza. Entretanto, o processo de avaliação de ajustes de diferentes funções usando o erro do ajuste como critério de mérito para a escolha de uma função em particular é vulnerável a diversas fontes de ruído.

Dados incompletos (o aparato experimental pode ser incapaz de detectar eventos para certas regiões do espectro), fontes de ruído externas ao fenômeno de interesse mas inerentes ao método de medida e o erro estatístico (particularmente em regiões do espectro de baixa contagem) irão afetar significativamente a escolha da função de parametrização dos dados. Em todos esses casos, pode-se citar a tendência a se escolher uma função de ajuste com um número alto de graus de liberdade, visto que o ajuste será melhor exatamente porque a função consegue descrever também o ruído da medida, podendo ser muito mais complexa do que a distribuição original do fenômeno de interesse.

Torna-se então especialmente relevante poder medir o grau de viés que é adicionado aos dados quando procura-se parametrizá-los. Em um problema com diversas soluções possíveis, como formas de ajuste de funções, espera-se que seja escolhida a solução que atende os dados disponíveis e que tenha o menor grau de comprometimento com relação aos dados indisponíveis. Em outras palavras, entre funções de distribuição que descrevem adequadamente os dados, aquela que possuir o maior grau de incerteza é a que carrega o mínimo de suposição com relação à informação desconhecida ou não acessível. Essa incerteza claramente é bem descrita pelo conceito de entropia da informação [7], visto que foi exatamente procurando qual grandeza reflete essa propriedade que Shannon desenvolveu o trabalho referido na seção anterior.

Isso é em essência o *método da entropia máxima* (MEM): entre todas as soluções possíveis, deve-se escolher aquela que tem a entropia máxima. O método foi desenvolvido por E. T. Jaynes em 1957, como parte do seu trabalho no desenvolvimento da *termodinâmica da entropia máxima* [1, 2], fundamentado na mecânica estatística de Gibbs, na teoria da

informação de Shannon, e em probabilidade Bayesiana.

Levando em consideração sua relação com a entropia física, uma solução com entropia máxima será a mais “natural”. Da perspectiva estatística e da teoria da informação, ela será a mais provável de todos os cenários devido a sua baixa complexidade. No sentido contrário, a solução com menor entropia seria excessivamente complexa e absolutamente improvável, dependente de uma série de mecanismos não evidenciados pelos dados e particular a poucos cenários que poderiam dar origem às medidas. Nesse sentido, é claro o paralelo entre o MEM e o princípio da *Navalha de Occam*, comumente enunciado como: “entre diferentes hipóteses, aquela com o menor número de suposições deve ser selecionada”.

Os argumentos acima fornecem então um procedimento para que, uma vez que se tenha em mãos diversas soluções que satisfazem o problema de interesse, seja possível selecionar qual solução é aquela que é mais fiel aos dados, no sentido de que ela seja a menos dependente de premissas externas às medições, e logo mais provável. Entretanto essa forma de usar o MEM ainda não resolve o problema de como encontrar estas soluções. Em problemas suficientemente complexos, o número de variáveis é excessivamente elevado, o que tradicionalmente levou a técnicas de simplificar o problema tratando isoladamente diferentes regiões do espectro de variáveis, ou levando em consideração apenas as variáveis mais significativas e desconsiderando as outras. A *análise linear de sistemas* é um exemplo claro dessa estratégia, simplificando a complexidade do problema para condições de contorno restritas tal que se possa tratar sistemas não lineares como se o fossem.

Como encontrar então soluções para problemas excessivamente complexos? Esta é uma pergunta que entra diretamente na área de *otimização matemática* e corresponde a um processo de busca no espaço de soluções. O presente trabalho não pretende resolver uma pergunta ampla como essa; escolhe-se o *método dos multiplicadores de Lagrange* para que as variáveis utilizadas definam a forma da solução e a parametrização da distribuição longitudinal. Essa solução será a forma da função de distribuição.

3.3 Método dos multiplicadores de Lagrange

Desenvolvido por J. L. Lagrange, o *método dos multiplicadores de Lagrange* [8] é uma das técnicas de otimização matemática mais utilizadas para encontrar extremos locais sujeitos a condições de contorno.

Problemas de otimização matemática são problemas em que o objetivo é encontrar

a melhor solução entre todas as soluções que satisfazem as condições de contorno do problema, em que o critério de mérito da solução é definido de maneira numérica, sendo uma função das variáveis de interesse chamada de função objetivo. O caso de interesse deste trabalho restringe-se às variáveis reais, não sendo de interesse as técnicas utilizadas para variáveis inteiras e de outros tipos.

Deseja-se sempre encontrar um extremo dessa função; seja um mínimo, como no caso de redução de consumo de recursos limitados, ou um máximo, como no caso de melhoria de uma medida de performance. Por essa razão, técnicas de otimização são frequentemente chamadas de técnicas de minimização – quando a quantidade de interesse deve ser maximizada, simplesmente toma-se como função objetivo o negativo da mesma. Neste trabalho o método de interesse é o de entropia **máxima**, tal que os termos *ótimo*, *máximo* e *extremo* irão sempre se referir ao valor desejado da função objetivo.

O problema de determinar os extremos de uma função corresponde a procurar máximos ou mínimos globais, para os quais geralmente não há uma estratégia analítica de encontrá-los. Além disso também não há, de maneira geral, uma forma de reconhecer se a solução encontrada é o extremo global da função objetivo ou meramente o extremo local da sua vizinhança. A menos que a estrutura do problema permita a existência de uma forma analítica de encontrar todos os extremos de interesse e esses extremos forem finitos, (ou tiverem um conjunto finito de formas analíticas) invariavelmente as técnicas de otimização recorrerão a estratégias exaustivas para varrer ou sondar o espaço de soluções.

Devido ao tamanho dos problemas tratados e os limites de busca de soluções em problemas numéricos, a grande parte dos métodos de otimização são algoritmos de busca local. O método dos multiplicadores de Lagrange faz parte deste grupo, mas antes de descrever como ele funciona, é necessário abordar como a busca local geralmente é realizada.

Na figura 3.3 é apresentado um exemplo simples de uma função objetivo a ser maximizada. Nesse caso a função possui dois máximos locais, sendo o pico em vermelho a vizinhança do máximo global. Em ambos os máximos, o gradiente da função será o vetor nulo, enquanto em todos os outros pontos o gradiente será um vetor que aponta na direção da maior taxa de crescimento da função.

É simples perceber que, aplicando-se um gradiente em todo o espaço de variáveis, será gerado um campo que aponta para os dois máximos de intensidade. Esse campo pode ser associado à uma força de atração: se um objeto é sujeito à ação do gradiente, ele convergirá para um dos dois picos, conforme a figura 3.4 onde é representado o gradiente da função objetivo.

As técnicas que utilizam o gradiente podem ser categorizadas como parte da família de

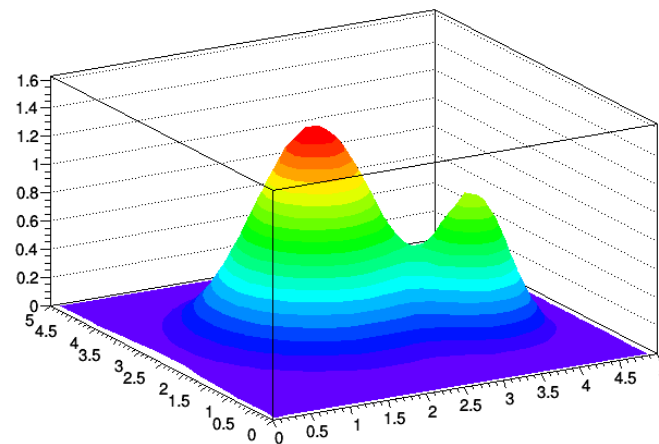


Figura 3.3: Exemplo de superfície de uma função objetivo de duas variáveis a ser maximizada, em que o valor da função é representado em cores e no eixo vertical do gráfico.

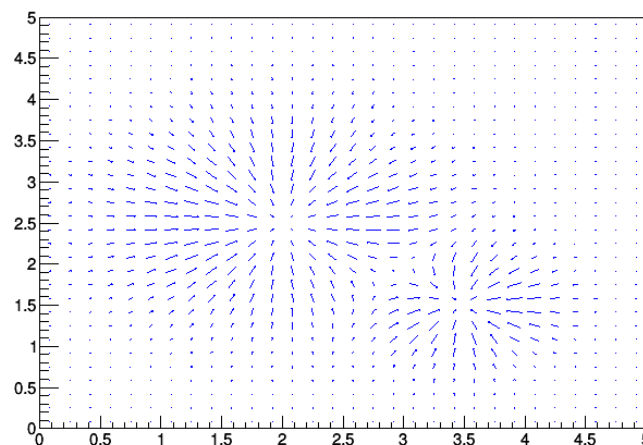


Figura 3.4: Representação da aplicação do gradiente ao exemplo de função objetivo da figura anterior; a orientação das setas indica o sentido do gradiente, e seu tamanho indica a magnitude.

métodos do gradiente. Esses métodos utilizam de maneiras variadas o gradiente da função objetivo, ou uma função de qual a função objetivo faz parte, para determinar qual direção seguir no processo de busca para encontrar o processo local. Assim o mapa de gradiente, como o do exemplo bidimensional da figura 3.4, serve como um indicador do caminho a ser seguido. Essa metodologia é análoga a abandonar uma pequena bola em um ponto do relevo: se a bola descer com velocidade suficientemente baixa, ela irá parar no fundo do vale que contém o ponto inicial do percurso.

Essa analogia é muito importante para entender o conceito de *bacia de atração*, que nada mais é que a região de soluções associada a um extremo local em que, se for utilizado

um método de gradiente, a busca convergirá para aquele extremo local. Retornando ao exemplo das figuras, existem duas bacias de atração com suas regiões bem determinadas no mapa de gradiente, bem definidas pelas duas áreas de convergência.

O método dos multiplicadores de Lagrange utiliza as condições de contorno, seus gradientes e o gradiente da função objetivo, tal que ele faz parte dessa categoria de métodos do gradiente. Ele foi desenvolvido para encontrar máximos ou mínimos locais sujeitos a condições de contorno do tipo igualdade (condições de contorno do tipo desigualdade são as *condições de Karush-Kuhn-Tucker* podem ser tratadas usando uma das generalizações do método de Lagrange, entretanto este não é o foco deste trabalho).

Seja um problema clássico de otimização restrita a obtenção do valor máximo da função objetivo f , com n variáveis e sujeita a m condições g_i :

$$\begin{aligned} \max \quad & f(\vec{x}) \\ & g_i(\vec{x}) = 0, \quad i \in \{1, \dots, m\} \end{aligned} \quad (3.9)$$

Cada condição de contorno define uma hipersuperfície (possivelmente degenerada) do espaço n -dimensional. A menos que as condições de contorno não possuam pontos em comum, o seu conjunto irá definir uma região que as satisfazem; no caso tridimensional, duas equações lineares de contorno com intercessão restringiriam o espaço de soluções a uma linha de soluções permissíveis, por exemplo.

Seja considerado então um ponto que pertence a essa região. Para transladar deste ponto para outro, sem sair da região permitida pelas condições de contorno, é necessário seguir em uma direção que seja ortogonal a cada hipersuperfície. Pode-se construir então um vetor ortogonal às direções permissíveis de deslocamento como a combinação linear dos gradientes das equações de contorno,

$$\sum_{i=1}^m \{\lambda_i \cdot \nabla_{\vec{x}} g_i(\vec{x})\} \quad (3.10)$$

onde λ_i são constantes.

Recuperando a ideia de que o gradiente da função objetivo $\nabla_{\vec{x}} f$ indica o caminho que leva ao máximo local da bacia de atração que contém o ponto avaliado, para todos os pontos em que $\nabla_{\vec{x}} f$ não for uma combinação linear dos m vetores $\nabla_{\vec{x}} g_i(\vec{x})$, há um ponto adjacente que é uma solução permissível e que possui um valor superior à posição avaliada; ou seja, a projeção de $\nabla_{\vec{x}} f$ sobre a região permissível é não nula, e pode-se seguir nesta direção para que se encontre uma solução superior que atende as condições de contorno.

Por outro lado, caso ocorra a igualdade

$$\nabla_{\vec{x}} f(\vec{x}) = \sum_{i=1}^m \{\lambda_i \cdot \nabla_{\vec{x}} g_i(\vec{x})\} \quad (3.11)$$

e $\nabla_{\vec{x}} f$ seja não nulo, então não existem pontos adjacentes que estejam dentro da região permissível e que tenham valor superior ao ponto avaliado: este é um ótimo local sob condições de contorno. Por outro lado, caso o $\nabla_{\vec{x}} f$ seja nulo, existem três possibilidades: \vec{x} é um mínimo, ou um ponto de sela, ou um máximo irrestrito pelas condições de contorno. Os dois primeiros casos são trivialmente checáveis avaliando-se a Hessiana da função objetivo no ponto, enquanto o terceiro não é um problema, visto que mesmo uma solução irrestrita pode fazer parte da região restrita pelas condições de contorno e que a equação 3.11 satisfaz $\nabla_{\vec{x}} f$ com valores nulos de λ_i .

Desta forma, um máximo local restrito pelas condições de contorno obedecerá as equações:

$$\nabla_{\vec{x}} f(\vec{x}) - \sum_{i=1}^m \{\lambda_i \cdot \nabla_{\vec{x}} g_i(\vec{x})\} = 0 \quad (3.12)$$

$$g_j(\vec{x}) = 0, \quad i \in \{1, \dots, m\} \quad (3.13)$$

Essas equações podem ser reescritas na forma da definição da *função de Lagrange* \mathcal{L}

$$\mathcal{L}(\vec{x}, \vec{\lambda}) = f(\vec{x}) - \sum_{i=1}^m \{\lambda_i \cdot g_i(\vec{x})\} \quad (3.14)$$

e utilizando o gradiente não só em \vec{x} , mas também em $\vec{\lambda}$

$$\nabla_{\vec{x}, \vec{\lambda}} \mathcal{L}(\vec{x}, \vec{\lambda}) = 0 \quad (3.15)$$

tal que o resultado da aplicação do gradiente com relação às variáveis \vec{x} e $\vec{\lambda}$ implicam, respectivamente, as equações 3.12 e 3.13.

Isso é equivalente a um processo de otimização irrestrita, cuja função objetivo é a função de Lagrange. A ausência de condições de contorno explícitas reduz a complexidade computacional e o tempo de processamento envolvidos.

3.4 Aplicação dos métodos na parametrização da distribuição longitudinal

O método dos multiplicadores de Lagrange fornece uma estratégia para buscar um máximo sob condições de contorno conforme o desejado, entretanto sua aplicação direta não sugere a função da distribuição na forma como enunciado na seção 3.3. No entanto é possível desenvolver as equações 3.14 e 3.15, utilizando as condições de contorno provenientes dos dados experimentais, para obter uma distribuição cuja forma é baseada na escolha

dos dados experimentais relevantes. Utilizando o método da entropia máxima, a função objetivo a ser maximizada será a entropia da distribuição.

Para encontrar então a função de distribuição desejada definem-se:

- x_i : Um valor de profundidade do perfil longitudinal \vec{x} em g/cm^2 , $i \in \mathbb{N}$. Elas são discretas visto que o perfil é representado por meio de histogramas e será tratado numericamente.
- $p(\vec{x})$: A função de distribuição das partículas em função da profundidade.
- $S(p(\vec{x}))$: A entropia da informação da distribuição na forma

$$S(p(\vec{x})) = - \sum_i \{p(x_i) \cdot \log(p(x_i))\} \quad (3.16)$$

- $h_k(\vec{x})$: Uma função da profundidade, podendo ter forma de potências, logaritmo, etc; ela será utilizada para obter momentos estatísticos nas equações de contorno.
- c_k : O valor mensurável do momento estatístico associado à função $h_k(\vec{x})$, tal que

$$\sum_i \{h_k(x_i) \cdot p(x_i)\} = c_k \quad (3.17)$$

- $g_k(\vec{x})$: A equação 3.17 pode ser reescrita como uma condição de contorno que será atendida quando seu resultado for nulo. Esta equação de contorno terá então a forma

$$g_k(\vec{x}) = \sum_i \{h_k(x_i) \cdot p(x_i)\} - c_k \quad (3.18)$$

Com essas definições pode-se escrever a função de Lagrange para o problema

$$\begin{aligned} \mathcal{L}(\vec{x}, \vec{\lambda}) &= S(p(\vec{x})) - \sum_k \{\lambda_k \cdot g_k(\vec{x})\} \\ &= - \sum_i \{p(x_i) \cdot \log(p(x_i))\} - \sum_k \left\{ \lambda_k \cdot \left[\sum_i \{h_k(x_i) \cdot p(x_i)\} - c_k \right] \right\} \end{aligned} \quad (3.19)$$

É necessário então que, para todas as funções $p(\vec{x})$, o valor de $S(p(\vec{x}))$ seja máximo. Ou seja, variando a forma de $p(\vec{x})$:

$$\begin{aligned} 0 &= \frac{\partial \mathcal{L}}{\partial p} = - \sum_i \{\log(p(x_i)) + 1\} - \sum_k \left\{ \lambda_k \cdot \left[\sum_i \{h_k(x_i)\} - 0 \right] \right\} \\ &= \sum_i \{-\log(p(x_i)) - 1 - \sum_k \{\lambda_k \cdot h_k(x_i)\}\} \end{aligned} \quad (3.20)$$

Como um ponto da função de distribuição não pode ser dependente de outro ponto

$$\frac{dp(x_i)}{dp(x_j)} = 0 \quad \forall i \neq j \quad (3.21)$$

tal que

$$\log(p(x_i)) = -1 - \sum_k \{\lambda_k \cdot h_k(x_i)\} \quad (3.22)$$

$$\Rightarrow p(x_i) = \exp(-1) \cdot \exp\left(-\sum_k \{\lambda_k \cdot h_k(x_i)\}\right) \quad (3.23)$$

Como $\exp(-1)$ na expressão acima é meramente um fator de escala, e sabe-se que para a função de distribuição de probabilidade $\sum_i \{p(x_i)\} = 1$, pode-se simplificar a expressão com um parâmetro de escala α :

$$p(x_i) = \alpha \cdot \exp\left(-\sum_k \{\lambda_k \cdot h_k(x_i)\}\right) \quad (3.24)$$

Adicionalmente, para o tratamento do perfil longitudinal são escolhidos como condições de contorno o primeiro momento de profundidade da distribuição (dado por $\sum_i \{x_i \cdot p(x_i)\}$) e o momento do logaritmo da profundidade (dado por $\sum_i \{\log(x_i) \cdot p(x_i)\}$). A escolha do primeiro momento de profundidade é fundamental, visto a importância da profundidade do máximo do número de partículas do perfil. A escolha do momento logaritmo tem considerações mais práticas: é pequeno o número de partículas em alta profundidade, além do que a chegada da cascata de partículas à superfície da Terra faz com que o final da distribuição seja experimentalmente não obtível; isso faz com que os momentos estatísticos de ordem n sofram flutuações em função da profundidade inicial do chuveiro atmosférico. Por outro lado, o logaritmo é menos sensível aos valores elevados de profundidade, tal que ele será menos afetado que outros momentos estatísticos, inclusive o primeiro momento de profundidade.

Vistas estas considerações sobre a forma da função e as condições de contorno escolhidas, a função de distribuição assume a forma

$$\begin{aligned} p(x_i) &= \alpha \cdot \exp(-\lambda_1 \cdot x_i - \lambda_2 \cdot \log(x_i)) \\ &= \alpha \cdot x_i^{-\lambda_2} \cdot \exp(-\lambda_1 \cdot x_i) \end{aligned} \quad (3.25)$$

Outro fator que não foi levado em consideração na dedução da forma da função acima é o desconhecimento da profundidade relativa ao início da distribuição, o que no estudo do perfil longitudinal de chuveiros atmosféricos corresponde ao ponto de colisão do primário. O método dos multiplicadores de Lagrange não prevê este tipo de parâmetro livre, tal que a função será alterada deliberadamente para que haja um grau de liberdade adicional que acomode distribuições cuja origem é desconhecida, e também para que esta metodologia possa ser usada na parametrização da profundidade do primário, definido como x_0 :

$$p(x_i) = \alpha \cdot (x_i - x_0)^{-\lambda_2} \cdot \exp(-\lambda_1 \cdot (x_i - x_0)) \quad (3.26)$$

É claro que essa função, se usada literalmente nas computações, permitiria valores negativos ou até mesmo complexos para a distribuição, o que é inadmissível; o parâmetro x_0 serve unicamente para que se possa “deslizar” a distribuição ao longo do eixo de profundidade, visto que a cascata não tem início antes da colisão inicial. Para evitar estes problemas, define-se adicionalmente que $p(x_i) = 0$ para quaisquer valores em que $x_i - x_0 < 0$.

Uma última simplificação da função de distribuição pode ser feita definindo um novo parâmetro a com valor $a = \alpha \cdot \exp(\lambda_1 \cdot x_0)$, substituindo λ_2 por b (denominado o parâmetro de forma), e λ_1 por c (denominado o parâmetro de escala), tal que a forma final utilizada neste trabalho para a função de distribuição do perfil longitudinal será

$$p(x_i) = a \cdot (x_i - x_0)^{-b} \cdot \exp(-c \cdot x_i) \quad (3.27)$$

Pode-se perceber que, não acidentalmente, a forma da equação definida é equivalente à *função de Gaisser-Hillas* [27]

$$N(X) = N_{max} \cdot \left(\frac{X - X_0}{X_{max} - X_0} \right)^{\frac{X_{max} - X_0}{\lambda}} \cdot \exp\left(\frac{X_{max} - X}{\lambda} \right) \quad (3.28)$$

onde X é a variável de profundidade, N_{max} é o número máximo observado de partículas na profundidade X_{max} , e X_0 e λ são parâmetros utilizados para inferir a composição química e energia do primário. Essa função foi proposta em 1977 e tradicionalmente utilizada na parametrização do perfil longitudinal de chuviros atmosféricos extensos.

3.5 Método do Lagrangiano aumentado

Uma forma de melhorar a convergência da otimização utilizada na obtenção da entropia máxima é utilizar não o próprio método dos multiplicadores de Lagrange apresentado na seção 3.3 (página 32), mas sim o método do Lagrangiano aumentado. Estudado por Powell, Hestenes e Rockafellar [28, 29], este método apresenta algumas similaridades aos métodos de penalidades da área de otimização matemática.

Os métodos de penalidades substituem o problema de otimização restrita (sujeita a condições de contorno) por um problema de otimização irrestrita ao adicionar à função objetivo f termos de penalidade que utilizam as próprias condições de contorno. Para objetivos de maximização sob condições de contorno do tipo igualdade, o problema pode então ser descrito na forma:

$$\max \quad f(\vec{x}) - \mu \cdot \sum_{i=1}^m \{g_i^2(\vec{x})\} \quad (3.29)$$

Assim pode-se iniciar as iterações com valores pequenos de μ para cada condição de contorno, que vão sendo levados ao infinito até que as restrições sejam suficientemente atendidas – isto é, g_i suficientemente próximo de zero para todas as m condições de contorno.

O método do Lagrangiano aumentado parte então da forma usual do método dos multiplicadores de Lagrange e adiciona um termo de penalidade semelhante ao descrito acima, podendo ser apresentado como o problema de maximização (que é o caso de interesse)

$$\max \quad f(\vec{x}) - \sum_{i=1}^m \{\lambda_i \cdot g_i(\vec{x})\} - \frac{\mu}{2} \cdot \sum_{i=1}^m \{g_i^2(\vec{x})\} \quad (3.30)$$

atualizando os estimadores dos multiplicadores de Lagrange a cada iteração k na forma

$$\lambda^{[k+1]} = \lambda^{[k]} + \mu^{[k]} \cdot g_i(\vec{x}) \quad (3.31)$$

O uso do método do Lagrangiano aumentado ao invés de um método puro de penalidades apresenta a vantagem de, ao utilizar estimadores dos multiplicadores de Lagrange, atender as condições de contorno sem ser necessário elevar o valor de μ ao infinito. Comparado ao método puro dos multiplicadores de Lagrange, o aumento do valor de μ em função da evolução dos valores da função objetivo e das condições de contorno permite convergência muito mais rápida para um máximo local que atende às restrições do problema.

3.6 Otimização numérica em programação não-linear com o ALGENCAN

ALGENCAN [9] é um programa desenvolvido por pesquisadores do Departamento de Matemática Aplicada do IMECC/Unicamp e pelo Departamento de Ciência da Computação do IME/USP com o propósito de obter soluções para problemas de programação não linear extremamente complexos em tempo de processamento moderado.

O algoritmo usa o método do Lagrangiano aumentado, sendo os subproblemas envolvidos resolvidos pelo GENCAN; esta biblioteca também é desenvolvida pelo mesmo grupo, utilizando a metodologia de resolução de problemas irrestritos para lidar com condições de contorno ativas. Como o programa não utiliza manipulações matriciais usuais na solução deste tipo de problema, ele é capaz de resolver problemas extremamente complexos em tempo de processamento moderado.

O algoritmo de solução de problemas propriamente dito é escrito em Fortran. Versões antigas do algoritmo eram escritas em Fortran 77, possuindo camadas de interface para

descrição do problema de interesse em linguagens mais modernas, como C/C++, Python, Ampl, Java, Tcl, e scripts em pseudo-linguagens utilizadas pelos ambientes CUTEr, Octave, Matlab, e R. A versão utilizada no presente trabalho é a 3.1.0, atualização mais recente de novembro de 2016; ela é escrita em Fortran 90, com interfaces disponíveis somente para C/C++, Ampl, Fortran 77, e CUTEst.

O interesse da ferramenta para este trabalho está na utilidade do ALGENCAN para resolver problemas de maximização de funções computacionalmente complexas devido a presença de potências e exponenciais no problema de otimização restrita de interesse.

4. RESULTADOS E DISCUSSÕES

4.1 Dados da simulação de chuviros com o CORSIKA

Para gerar os dados de partículas do chuviro atmosférico, que foram tratados com o método da entropia máxima, foi utilizado o programa de simulação CORSIKA, conforme descrito em 2.4 (página 26). Apresenta-se na figura 4.1 os gráficos do número de partícula em função da profundidade em $g \cdot cm^{-2}$, para cada tipo de partícula, a partir dos dados de uma simulação de chuviro atmosférico com primário de $10^{17} eV$.

Na figura 4.2 são comparados os perfis de dois chuviros simulados sob as mesmas condições mas com colisões primárias em profundidades diferentes. Os gráficos apresentam o número de partículas por intervalo de profundidade, sendo a profundidade contada em intervalos de $10g \cdot cm^{-2}$. Para a primeira distribuição a colisão do primeiro ocorre no topo da atmosfera, tal que o número de partículas carregadas no intervalo entre $5 \pm 5g \cdot cm^{-2}$ é muito superior a uma partícula (a primária). A colisão inicial da segunda distribuição pode ser considerada profunda, de forma que conta-se apenas a partícula primária em profundidades inferiores a da primeira colisão que ocorre dentro do intervalo $155 \pm 5g \cdot cm^{-2}$. Fica claro que os perfis são muito semelhantes apesar da simulação pseudo aleatória de cada evento de interação das partículas; a maior diferença entre as duas é o *shift* horizontal devido à diferença entre as profundidades iniciais.

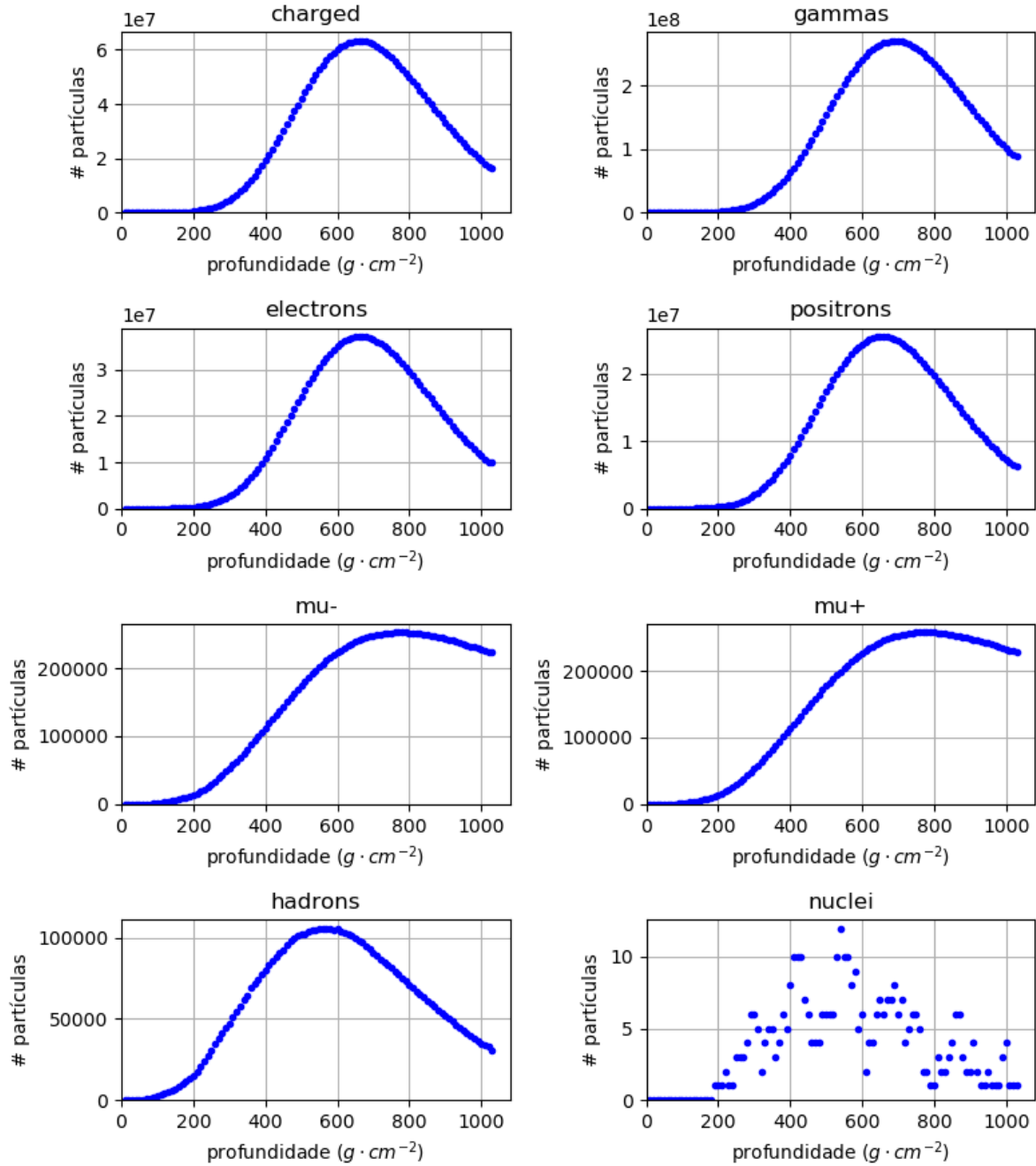


Figura 4.1: Exemplo de dados de output da simulação do chuveiro extenso com um próton de $10^{17}eV$ como primário gerado com o CORSIKA [3], apresentando o perfil longitudinal do número de diferentes espécies de partículas a cada $10g \cdot cm^{-2}$ de profundidade do chuveiro.

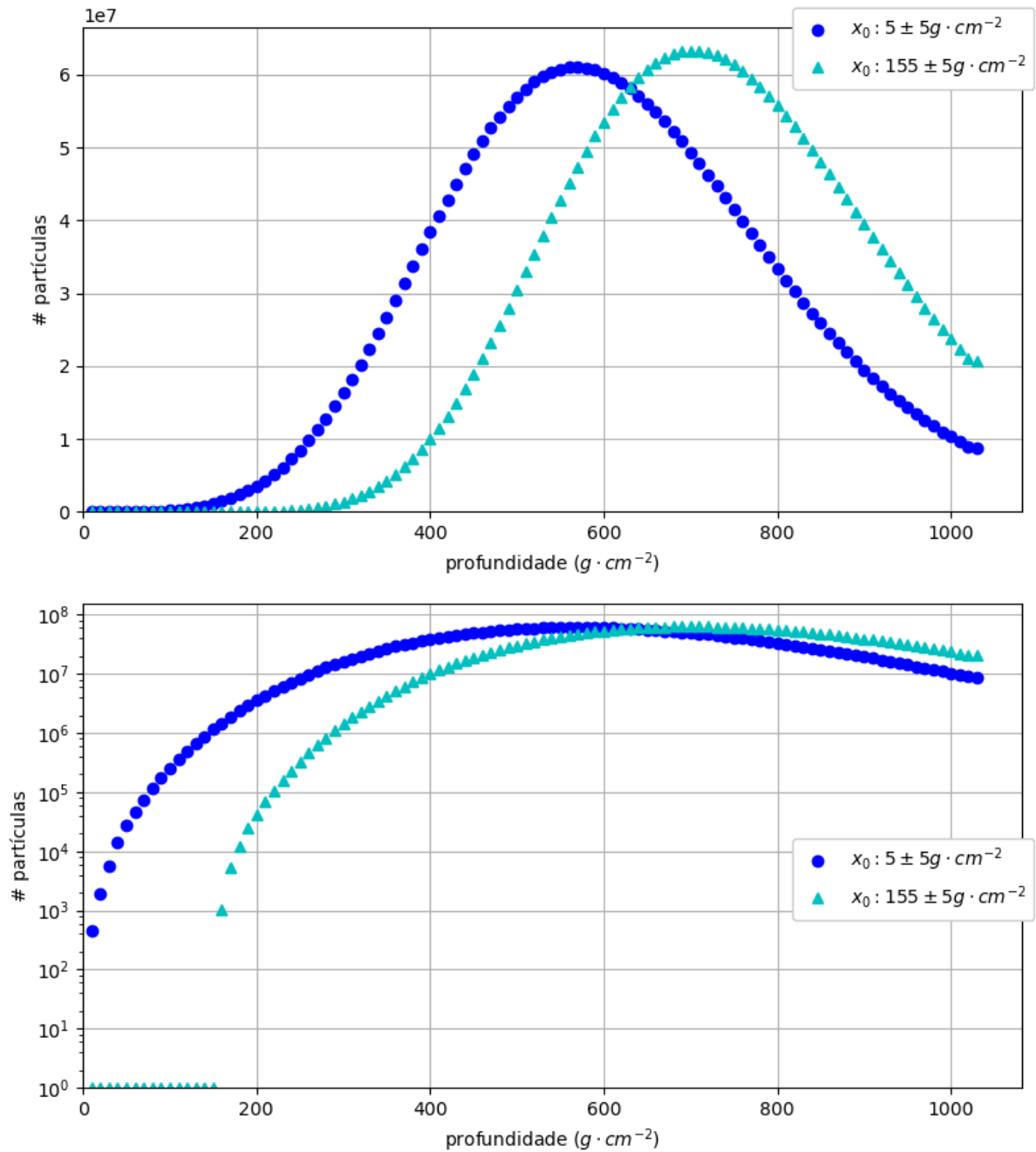


Figura 4.2: Comparação em escala linear e escala logarítmica entre os perfis longitudinais de partículas carregadas de duas simulações do CORSIKA [3] com as mesmas condições iniciais, sendo o primário um próton de $10^{17}eV$; a primeira simulação gerou uma colisão inicial a uma profundidade de $5 \pm 5g \cdot cm^{-2}$, enquanto na segunda ela ocorreu em $155 \pm 5g \cdot cm^{-2}$.

4.2 Aplicação do MEM para múltiplas distribuições vinculadas, e estudo dos perfis eletrônico e muônico de número e energia

Uma das grandes vantagens do método da entropia máxima é a possibilidade de se parametrizar com facilidade diversas distribuições diferentes em um único processo. Cada distribuição pode ter parâmetros diferentes ou vinculados, assim como suas condições de contorno, sendo então calculada a entropia particular de cada uma. É possível então maximizar a entropia total das distribuições consideradas atendendo as condições de contorno individuais, tal que as distribuições são parametrizadas conjuntamente em um único processo de otimização.

A vantagem da parametrização conjunta em único processo de otimização pertence, é claro, a casos de estudo em que existem parâmetros comuns às diferentes distribuições. Para o caso do estudo de perfis longitudinais, são apresentadas na figura 4.3 as distribuições de número de partículas e energia de ionização em função da profundidade para partículas eletrônicas e muônicas carregadas. O interesse de se parametrizar essas distribuições em conjunto vem do parâmetro de profundidade inicial da colisão ser um só para os quatro perfis, apesar de os mesmos tratarem de grandezas diferentes e espécies de partículas diferentes.

Uma dificuldade enfrentada em outras formas de ajustes de curva é qual critério utilizar para ponderar a minimização das curvas aos grupos de dados, visto que essa escolha pode polarizar o processo de otimização favorecendo o ajuste em uma curva com prejuízo ao ajuste de outra. Essa adaptação da parametrização não é necessária aqui porque as entropias individuais são função das probabilidades de cada intervalo, e não de valores absolutos; assim é possível aplicar o método da máxima entropia da mesma forma para uma distribuição como para várias distribuições diferentes que compartilhem parâmetros comuns.

Apesar do potencial dessa técnica, percebe-se na figura que a parametrização deixa a desejar para as distribuições de partículas muônicas carregadas. Isso ocorre porque a distribuição de Gaisser-Hillas [27], em sua versão original, descreve melhor alguns tipos de partículas que outros, não fornecendo um ajuste ideal para o perfil muônico. Não só isso, como essa escolha também interfere na parametrização da profundidade primária compartilhada pelos quatro perfis, causando resultados inferiores para a parametrização dos perfis eletrônicos.

Dessa forma, apesar de ter sido verificado o potencial de utilização do MEM para perfis

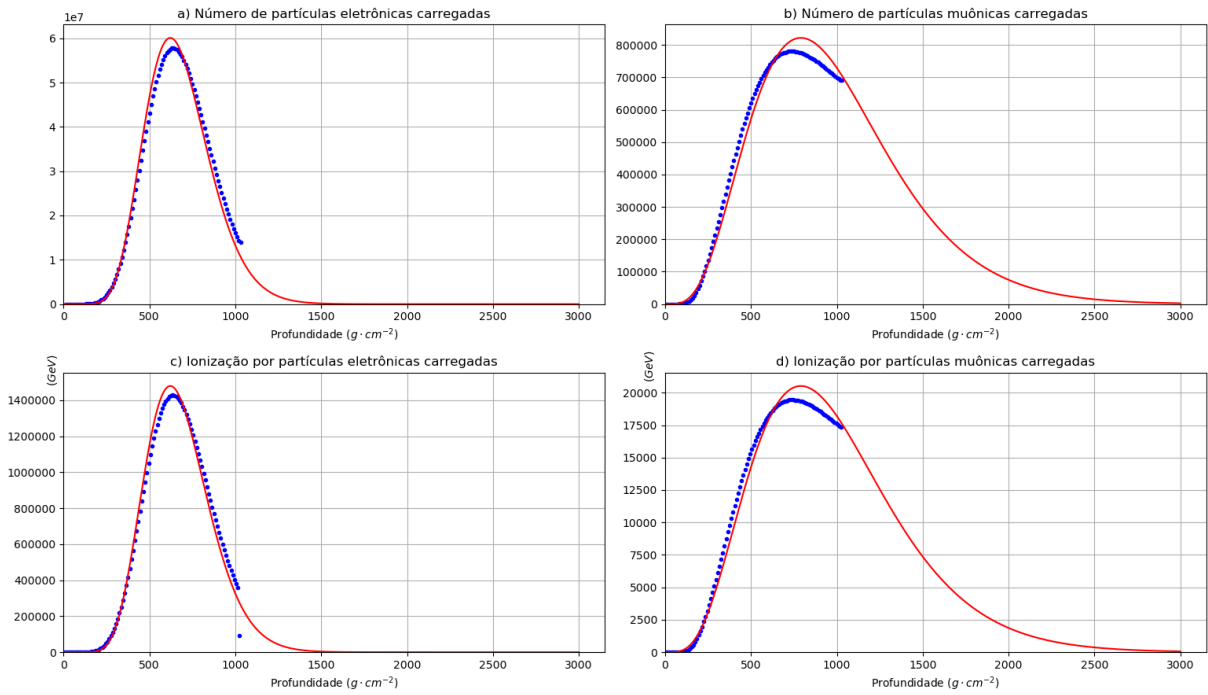


Figura 4.3: Aplicação do método da entropia máxima à entropia total das distribuições por profundidade de: número de partículas eletrônicas carregadas (a), número de partículas muônicas carregadas (b), energia de ionização por partículas eletrônicas carregadas (c), e energia de ionização por partículas muônicas carregadas (d). Os pontos azuis correspondem aos dados, e as curvas vermelhas representam as distribuições obtidas.

diferentes, foi feita a escolha de se trabalhar apenas com perfis eletrônicos carregados ou de partículas carregadas (predominantemente elétrons e positrons, como percebe-se na figura 4.1) deste ponto em diante.

4.3 Aplicação do MEM fora da região dos dados conhecidos

Outra vantagem do MEM é a possibilidade de utilizar a técnica mesmo em regiões de dados indisponíveis. Isso é vantajoso em situações em que parte do espaço de variáveis é não avaliável devido às condições experimentais, o número de eventos nessas regiões é muito pequeno, deseja-se excluir certos tipos de dados devido a *outliers*, ou a relação sinal ruído é muito baixa.

Como as condições de contorno devem ser aplicadas sobre a região dos dados conhecidos, mas a maximização da entropia pode ser realizada sobre toda a extensão desejada, o uso do conceito de entropia no MEM permite que mesmo regiões sem dados contribuam para a parametrização do comportamento avaliado.

Isso é fundamental para que as curvas obtidas não sejam viciadas pelo ruído inerente aos dados conhecidos, resultando em uma parametrização que seja a mais parcimoniosa e ainda respeite as condições de contorno, como descrito em 3.2 (página 30). Por outro lado, a redução da incerteza leva a distribuições que sejam as mais uniformes dentre as que respeitam as condições de contorno (3.1, página 27).

Um exemplo deste fenômeno é apresentado na figura 4.4. Neste exemplo, os dados disponíveis vão somente até uma profundidade em torno de $1000g \cdot cm^{-2}$, e a entropia foi maximizada até uma profundidade três vezes maior. A distribuição de Gaisser-Hillas permite que as condições de contorno de média linear e média logarítmica do número de partículas sejam atendidas por uma parametrização claramente inadequada aos dados. É fácil perceber também que, se ambas as distribuições atendem equivalentemente as condições de contorno, a distribuição obtida pela parametrização é a mais uniforme. No processo de maximização de entropia, o início da curva é achatado para que a longa cauda seja tão alta quanto seja possível respeitando as médias linear e logarítmica.

Essas condições de contorno são obviamente insuficientes para que a parametrização obtida represente adequadamente os dados. Isso orientou o rumo do trabalho então a uma avaliação da razão pela qual as médias utilizadas são insuficientes, e que outras condições de contorno seriam adequadas para obter uma parametrização fidedigna. Esse estudo é apresentado na seção a seguir.

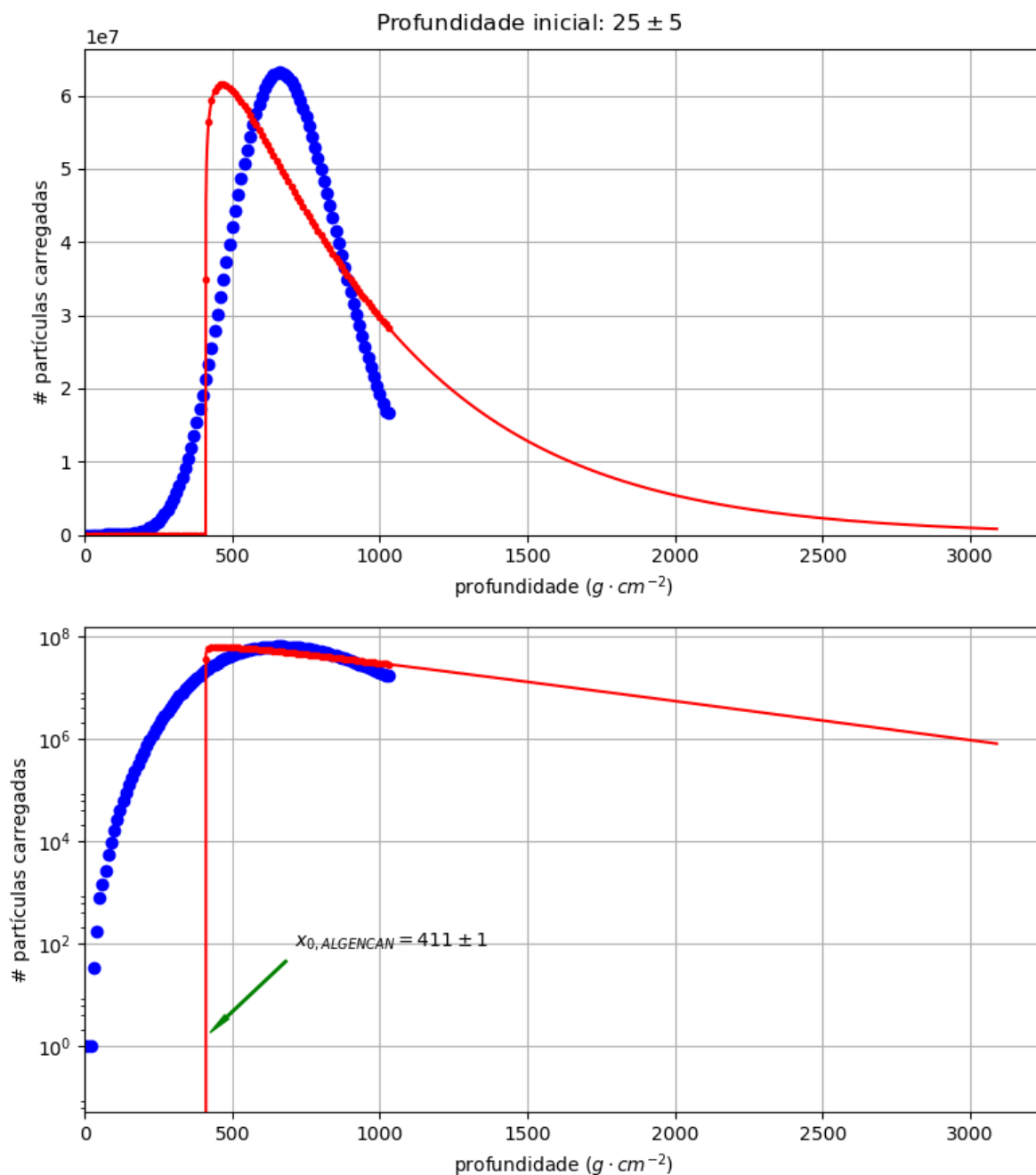


Figura 4.4: Resultado da aplicação do MEM em dados com profundidade até aproximadamente $1000 \text{ g} \cdot \text{cm}^{-2}$ considerando a entropia da distribuição parametrizada até aproximadamente $3000 \text{ g} \cdot \text{cm}^{-2}$, com apenas as condições de contorno de média linear e média logarítmica. Os dados são apresentados em azul, e a distribuição obtida é representada pela curva vermelha, com pontos nas posições correspondentes às quais há a presença de dados.

4.4 Estudo de máximos locais da entropia a partir dos parâmetros de forma e escala

Na figura 4.5 são apresentadas três distribuições com o mesmo parâmetro de forma b e três valores diferentes para o parâmetro de escala c , como descrito em 3.4 (página 36). Depreende-se que a mudança do parâmetro não afeta significativamente só a moda e a média linear da distribuição, mas ela causa uma mudança profunda na inclinação em escala logarítmica da cauda, visto que o parâmetro de escala é dominante no comportamento da curva para valores elevados de profundidade.

Em contraste, a figura 4.6 apresenta três distribuições agora com o parâmetro de escala c fixo, com três valores diferentes para o parâmetro de forma b . O parâmetro de forma afeta significativamente a inclinação em escala logarítmica da função para valores baixos de profundidade, afetando bastante o crescimento da distribuição. Por outro lado, em profundidade elevada as caudas tem inclinações logarítmicas extremamente semelhantes, o que condiz com a análise do parâmetro de escala.

Uma inferência importante, mas talvez não imediatamente óbvia, que pode ser feita a partir dos gráficos com eixo vertical em escala linear e logarítmica é que a média linear será muito mais sensível a mudanças nos parâmetros mencionados do que a média logarítmica. Para confirmar essa análise, foram criados os gráficos da figura 4.7, em que variam-se os valores do parâmetro de forma b no eixo horizontal e do parâmetro de escala no eixo vertical.

Os resultados são apresentados em escala cromática, indo do azul para baixas intensidades até o vermelho para altas intensidades: à esquerda são apresentados os valores da média logarítmica e à direita são apresentados os valores da média linear; os gráficos superiores mostram os valores absolutos, enquanto nos gráficos inferiores é feita a comparação da diferença relativa entre os valores absolutos obtidos da função e os valores fixos um caso de teste dos dados de simulação. Assim a região de parâmetros desejada, que melhor atende aos dados, aparecerá nos gráficos comparativos inferiores da figura 4.7 como regiões azul escuras.

Como inferido, a variação da média linear (à direita) é muito superior à da média logarítmica, representado como uma mudança muito mais rápida de cores ao longo do espaço bidimensional de parâmetros. Enquanto a região que atende à média linear em até 20% de variação é bastante restrita, quase toda a região exibida atende à média logarítmica com menos de 10% de variação. Não só isso, mas percebe-se que os critérios

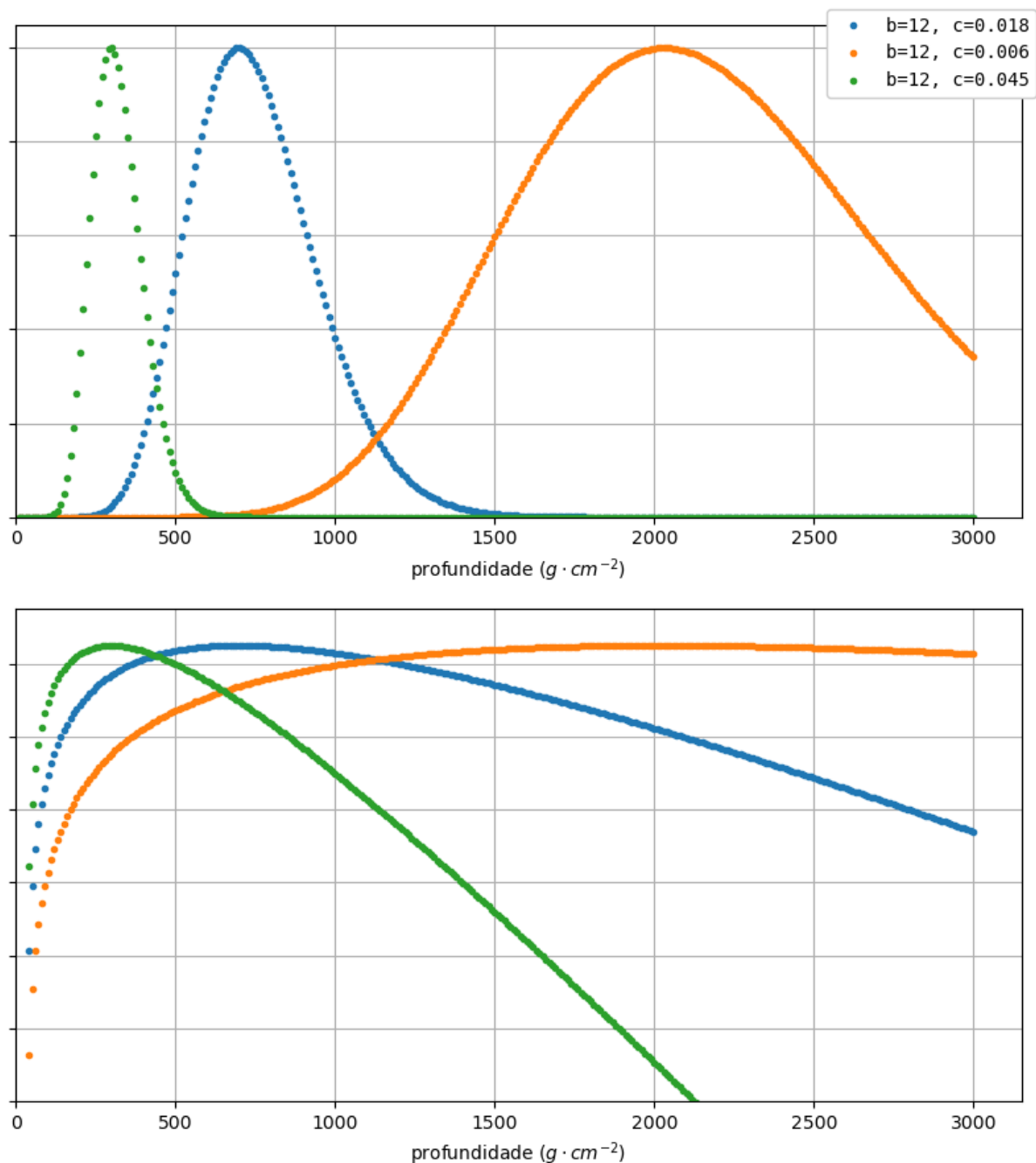


Figura 4.5: Distribuição de Gaisser-Hillas para um valor fixo do parâmetro de forma b , diferentes valores do parâmetro de escala c , e uma profundidade primária de $30 \text{g} \cdot \text{cm}^{-2}$. A figura superior tem o eixo y em escala linear, enquanto a inferior em escala logarítmica. Este eixo não apresenta valores em ambas escalas, visto que as três curvas sofreram diferentes normalizações para que seus máximos fossem iguais, meramente com o propósito de facilitar a comparação visual.

das médias linear e logarítmica são quase redundantes, sendo suas regiões de variação mínima praticamente as mesmas (a faixa escura diagonal estreita que atravessa os dois gráficos).

Para demonstrar mais claramente o que as regiões de menor variação da figura 4.7

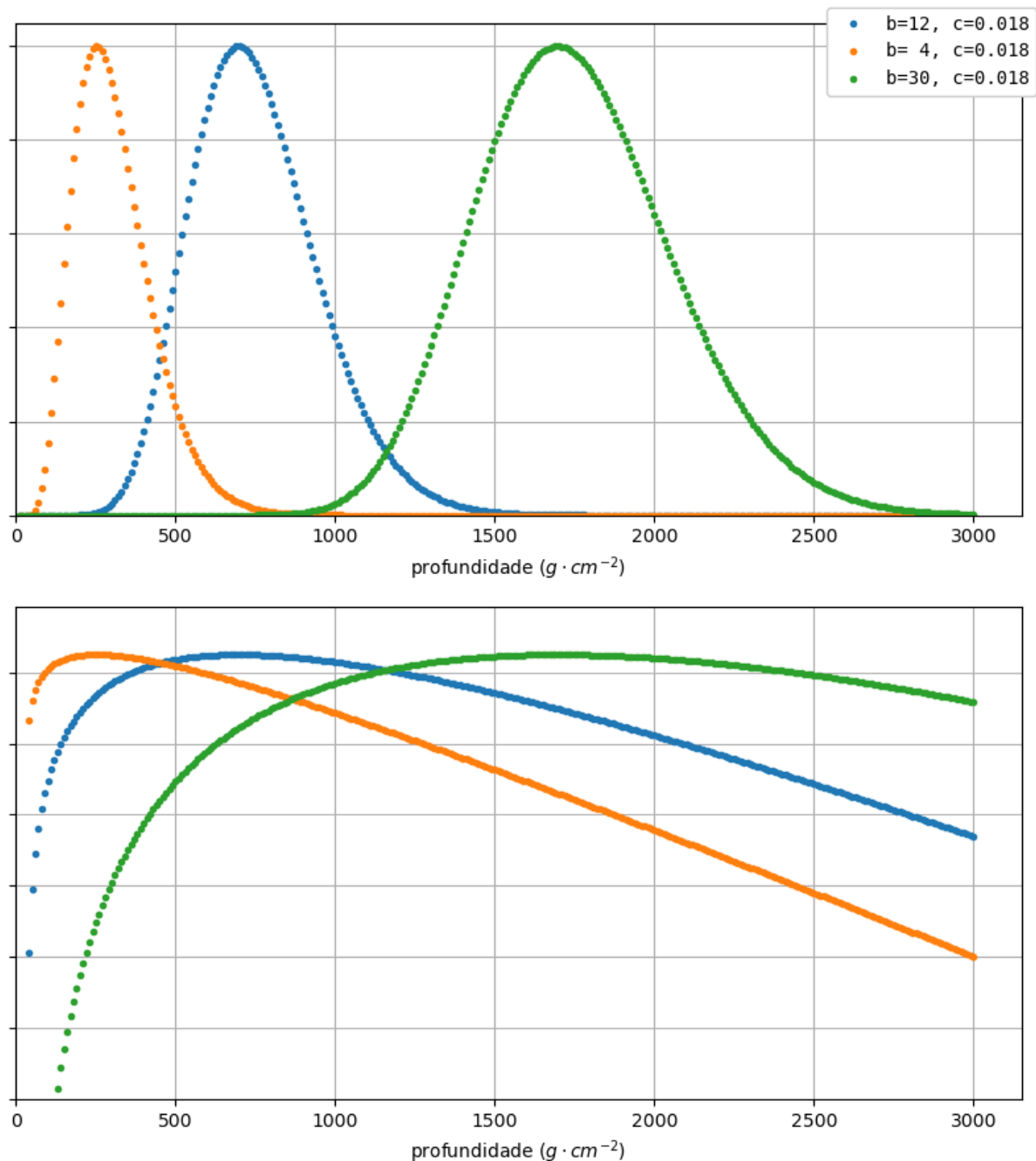


Figura 4.6: Distribuição de Gaisser-Hillas para diferentes valores do parâmetro de forma b , um valor fixo do parâmetro de escala c , e uma profundidade primária de $30g \cdot cm^{-2}$. A figura superior tem o eixo y em escala linear, enquanto a inferior em escala logarítmica. Este eixo não apresenta valores em ambas escalas, visto que as três curvas sofreram diferentes normalizações para que seus máximos fossem iguais, meramente com o propósito de facilitar a comparação visual.

significa, apresenta-se na figura 4.8 três distribuições com parâmetros de forma e escala ajustados de forma que todas atendam as condições de contorno.

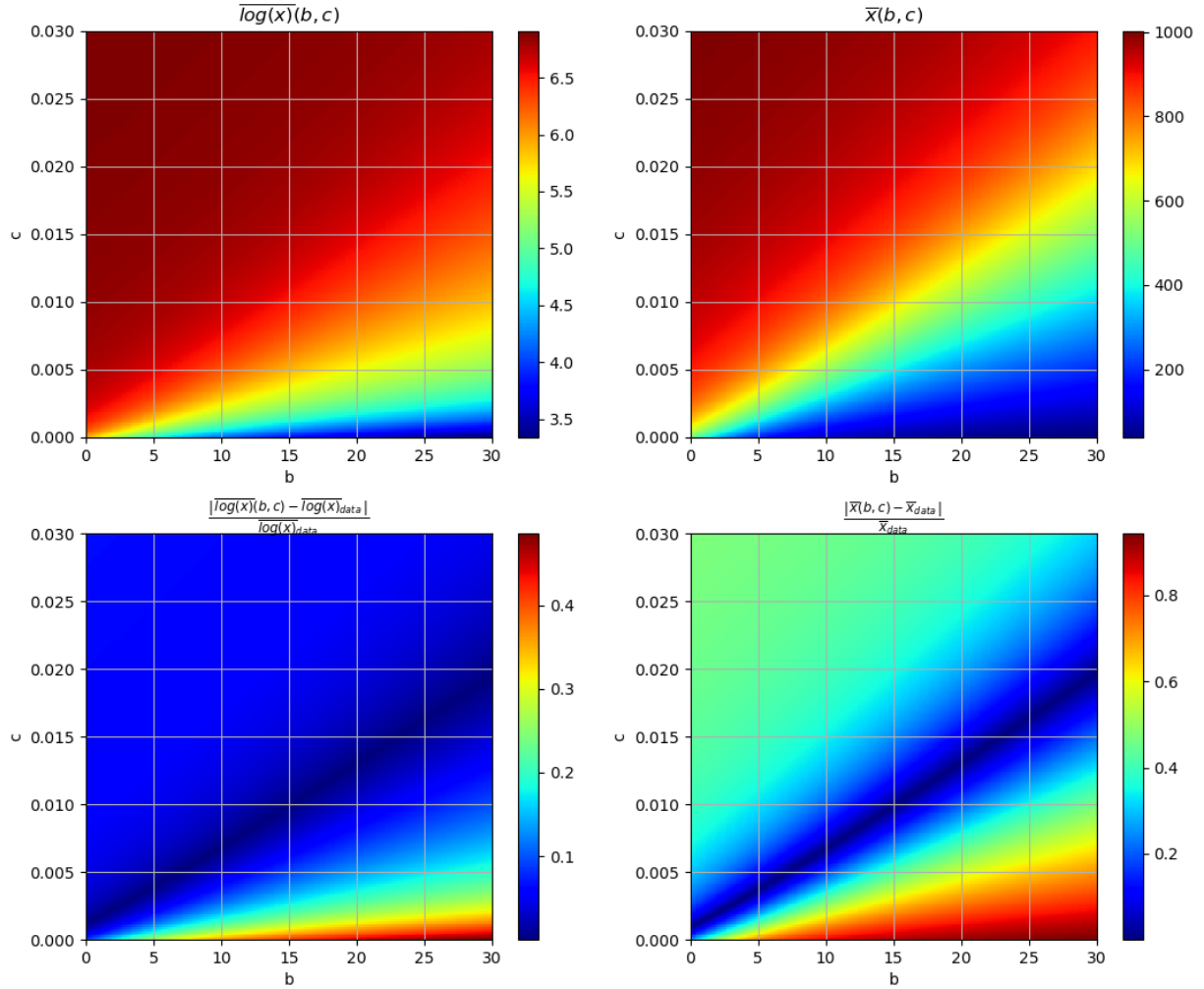


Figura 4.7: Gráficos bidimensionais de estatísticas obtidas a partir da distribuição de Gaisser-Hillas em função dos parâmetros de forma b , de escala c , e uma profundidade inicial nula, comparados aos dados de um perfil longitudinal simulados com o CORSIKA [3]. São representados a média do logaritmo da profundidade (superior esquerda), o módulo do desvio relativo desta média logarítmica com relação aos dados com $\log(x)_{data} \approx 6.5$ (inferior esquerda), a média da profundidade (superior direita), e o módulo do desvio relativo desta média com relação aos dados com $\bar{x}_{data} \approx 680$ (inferior direita). Todos os valores de profundidade em $g \cdot cm^{-2}$.

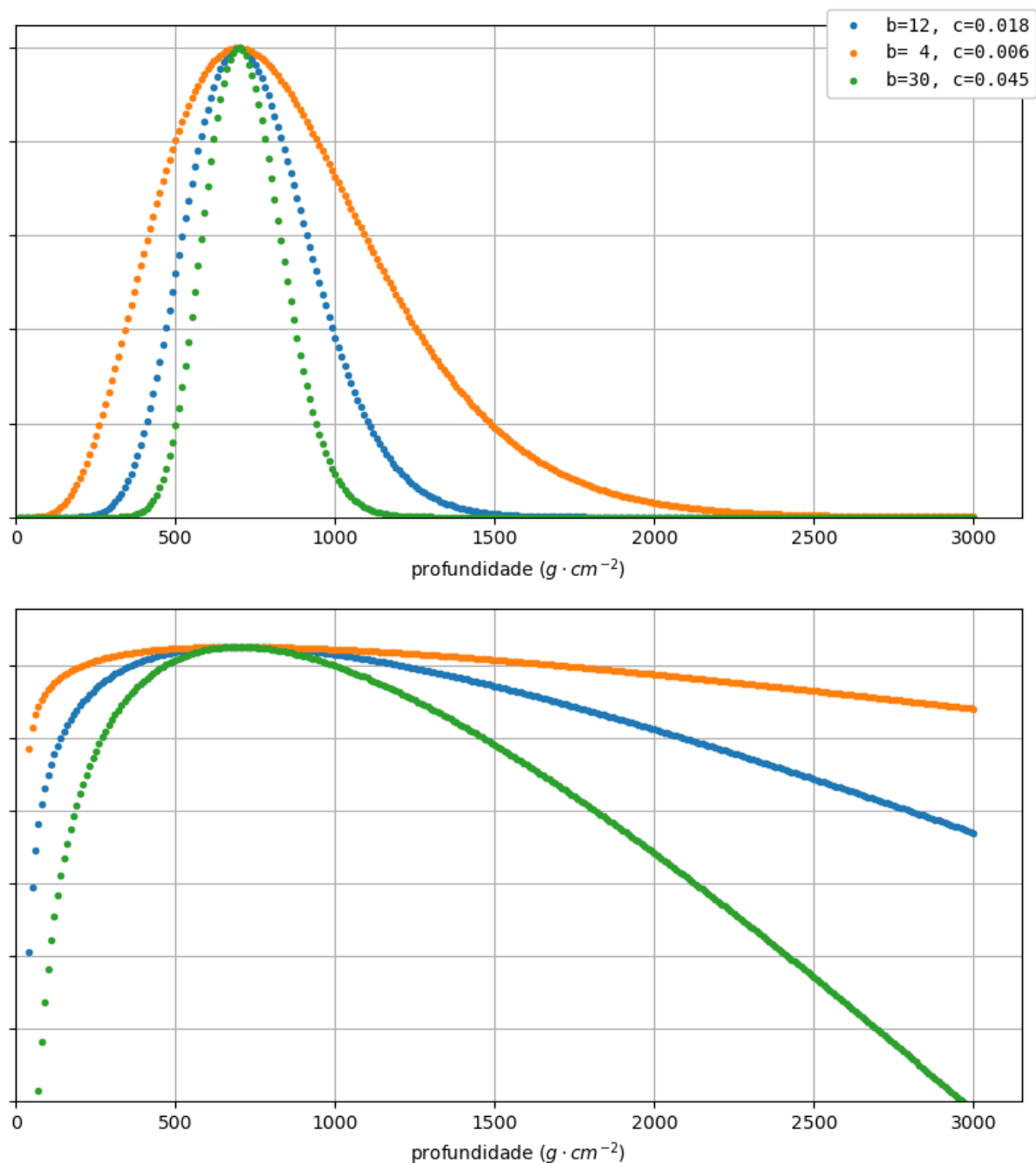


Figura 4.8: Distribuição de Gaisser-Hillas para diferentes valores dos parâmetros de forma b , do parâmetro de escala c , e uma profundidade primária de $30g \cdot cm^{-2}$. Os parâmetros b e c foram ajustados para que todas distribuições atendam as condições de contorno. A figura superior tem o eixo y em escala linear, enquanto a inferior em escala logarítmica. Este eixo não apresenta valores em ambas escalas, visto que as três curvas sofreram diferentes normalizações para que seus máximos fossem iguais, meramente com o propósito de facilitar a comparação visual.

Percebe-se que as distribuições resultantes são bastante diferentes apesar de atender satisfatoriamente as restrições, perceptivelmente redundantes para o intervalo de profundidade avaliado. Essa redundância não ocorreria caso o intervalo de profundidade fosse muito superior; no entanto é importante lembrar que os dados obtidos experimentalmente de chuviros atmosféricos são limitados pela profundidade da atmosfera, que não será muito superior a $1000g \cdot cm^{-2}$. Assim é inevitável que, dadas apenas as condições de contorno de média linear e logarítmica, seja sempre possível compensar os valores de um parâmetro alterando-se o outro. Em termos matemáticos, existe um número infinito de parâmetros de forma e escala que podem atender a ambas condições de contorno.

Utilizando o critério de entropia máxima, a distribuição resultante será aquela mais larga, por ser a mais uniforme e apresentar a maior incerteza. O reconhecimento deste problema contém invariavelmente a sua própria solução: é necessário limitar quão larga será a distribuição obtida pelo MEM. Basta então introduzir uma nova condição de contorno que seja estatisticamente resistente a ruído e facilmente mensurável como as médias utilizadas até então – e visivelmente a variância parece ser uma ótima candidata.

Antes de refazer a otimização com a nova condição de contorno, é razoável repetir a análise do espaço de soluções para a variância; como a média linear se mostrou mais sensível a mudanças nos parâmetros, ela é novamente apresentada à direita da avaliação da variância na figura 4.9. Novamente os pontos azul escuros dos gráficos inferiores indicam a menor variação entre os valores obtidos da função para os parâmetros e o valor desejado obtido dos dados. A região do espaço de parâmetros que melhor atende à variância dos dados não é uma faixa reta como para as médias linear e logarítmica, mas forma uma espécie de laço – essas regiões se interceptam, mostrando que a adição da condição da variância como condição de contorno adicional efetivamente resolve o problema do espaço infinito de soluções.

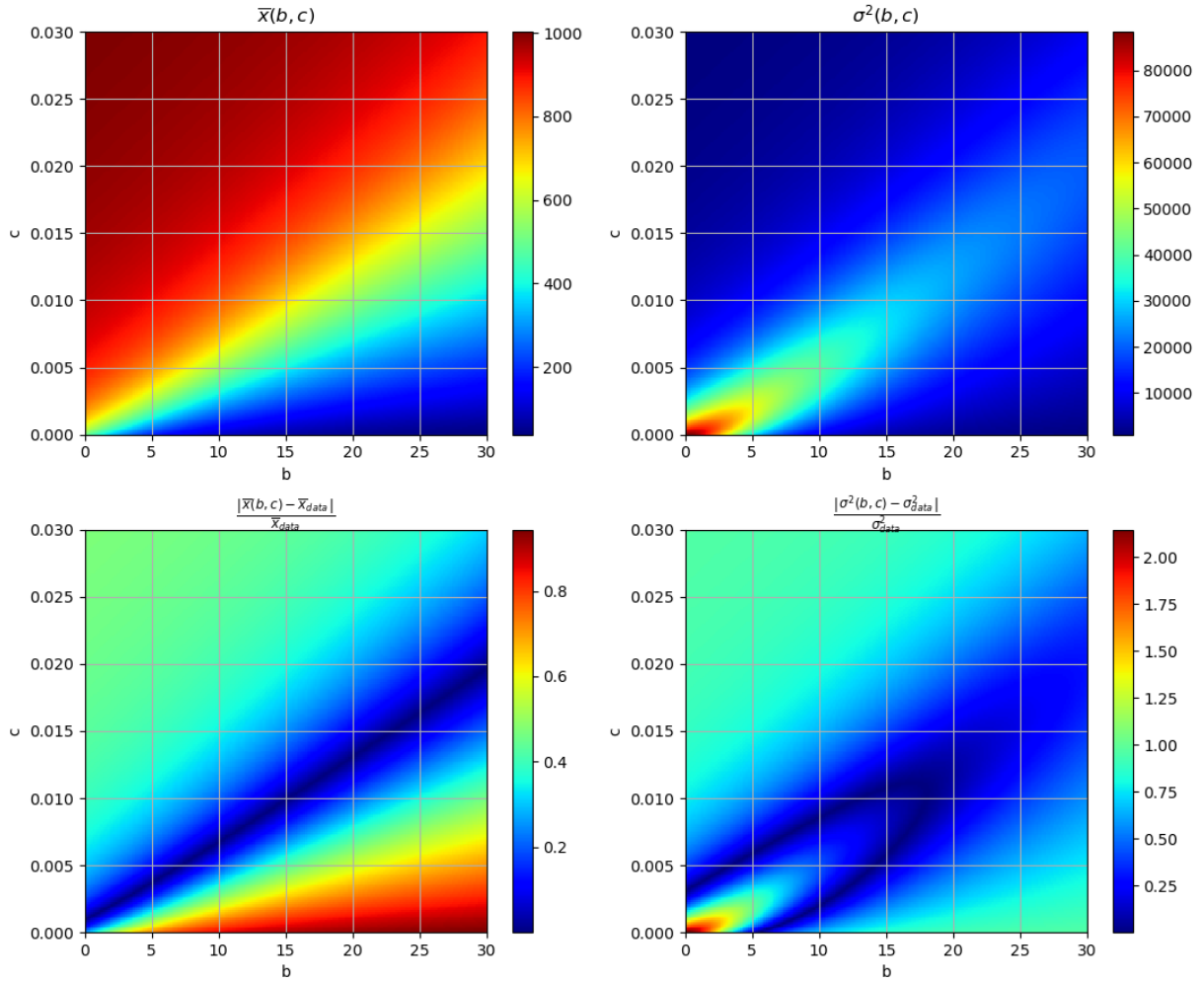


Figura 4.9: Gráficos bidimensionais de estatísticas obtidas a partir da distribuição de Gaisser-Hillas em função dos parâmetros de forma b , de escala c , e uma profundidade inicial nula, comparados aos dados de um perfil longitudinal simulados com o CORSIKA [3]. São representados a média da profundidade (superior esquerda), e o módulo do desvio relativo desta média com relação aos dados com $\bar{x}_{data} \approx 680$ (inferior esquerda), a variância da profundidade (superior direita), e o módulo do desvio relativo desta variância com relação aos dados com $\sigma^2_{data} \approx 28 \cdot 10^3$ (inferior direita). Todos os valores de profundidade em $g \cdot cm^{-2}$.

4.5 Introdução da terceira condição de contorno

Feita a confirmação da seção anterior, pode-se proceder a reimplementação do MEM com a adição da variância como condição de contorno, mantendo o critério de entropia máxima para um intervalo largo de profundidade (cerca de três vezes superior ao intervalo em que há dados experimentais).

São apresentados dois resultados para profundidades iniciais diferentes nas figuras 4.10 e 4.11.

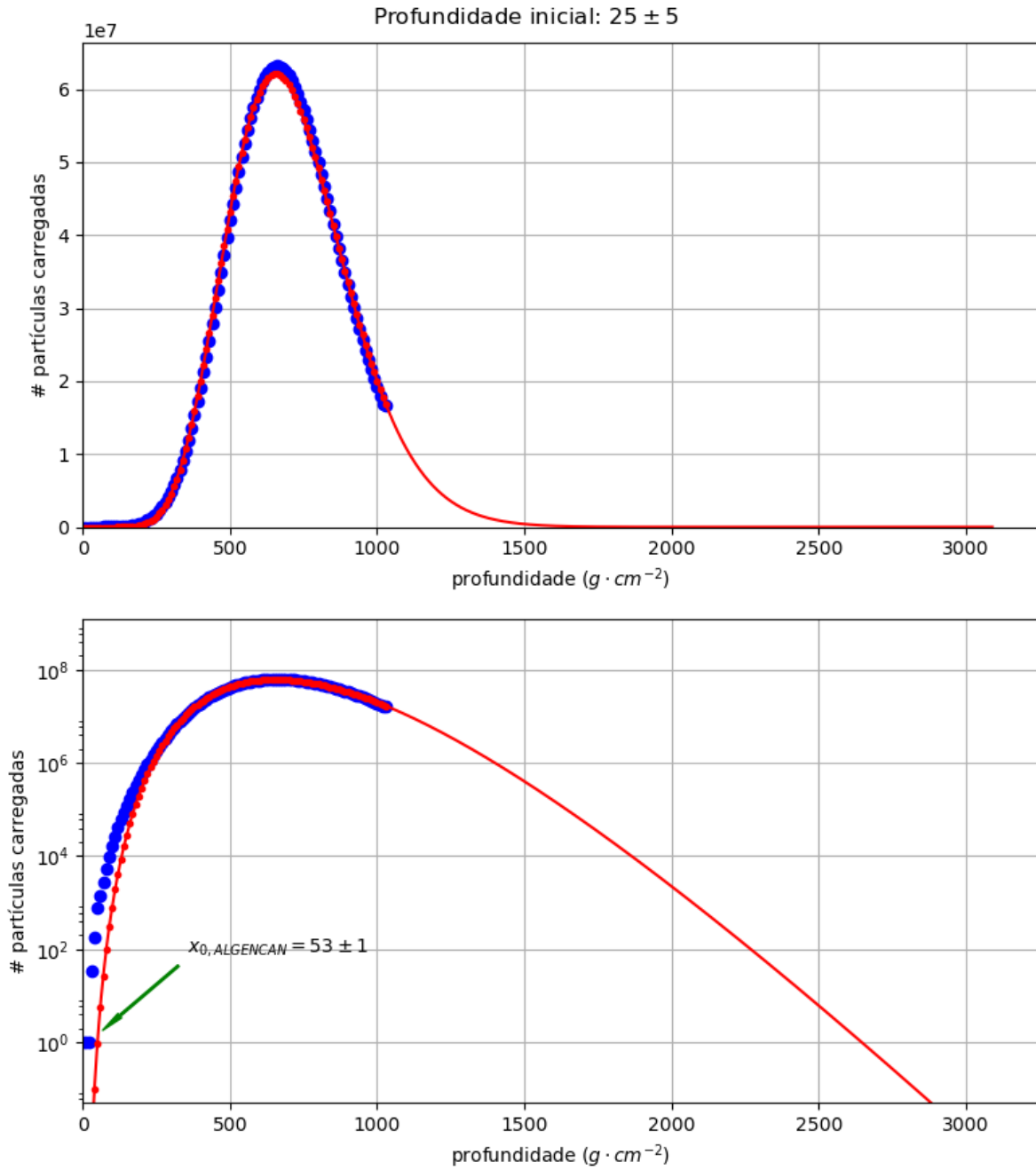


Figura 4.10: Resultado da aplicação do MEM em dados de colisão primária “rasa”, com simulação da cascata de partículas até uma profundidade de aproximadamente $1000 \text{g} \cdot \text{cm}^{-2}$, considerando a entropia da distribuição parametrizada até aproximadamente $3000 \text{g} \cdot \text{cm}^{-2}$, com as condições de contorno de média linear, média logarítmica, e variância dos dados. Os dados são apresentados em azul, e a distribuição obtida é representada pela curva vermelha, com pontos nas posições correspondentes às quais há a presença de dados.

Pelos gráficos com escala vertical linear, percebe-se que agora a parametrização é muito mais fiel à distribuição dos dados, não sofrendo a divergência que ocorria antes da última alteração. Por outro lado, o gráfico em escala logarítmica da figura 4.10 mostra que ainda há margem para avanços de parametrização na região do início da cascata de partículas.

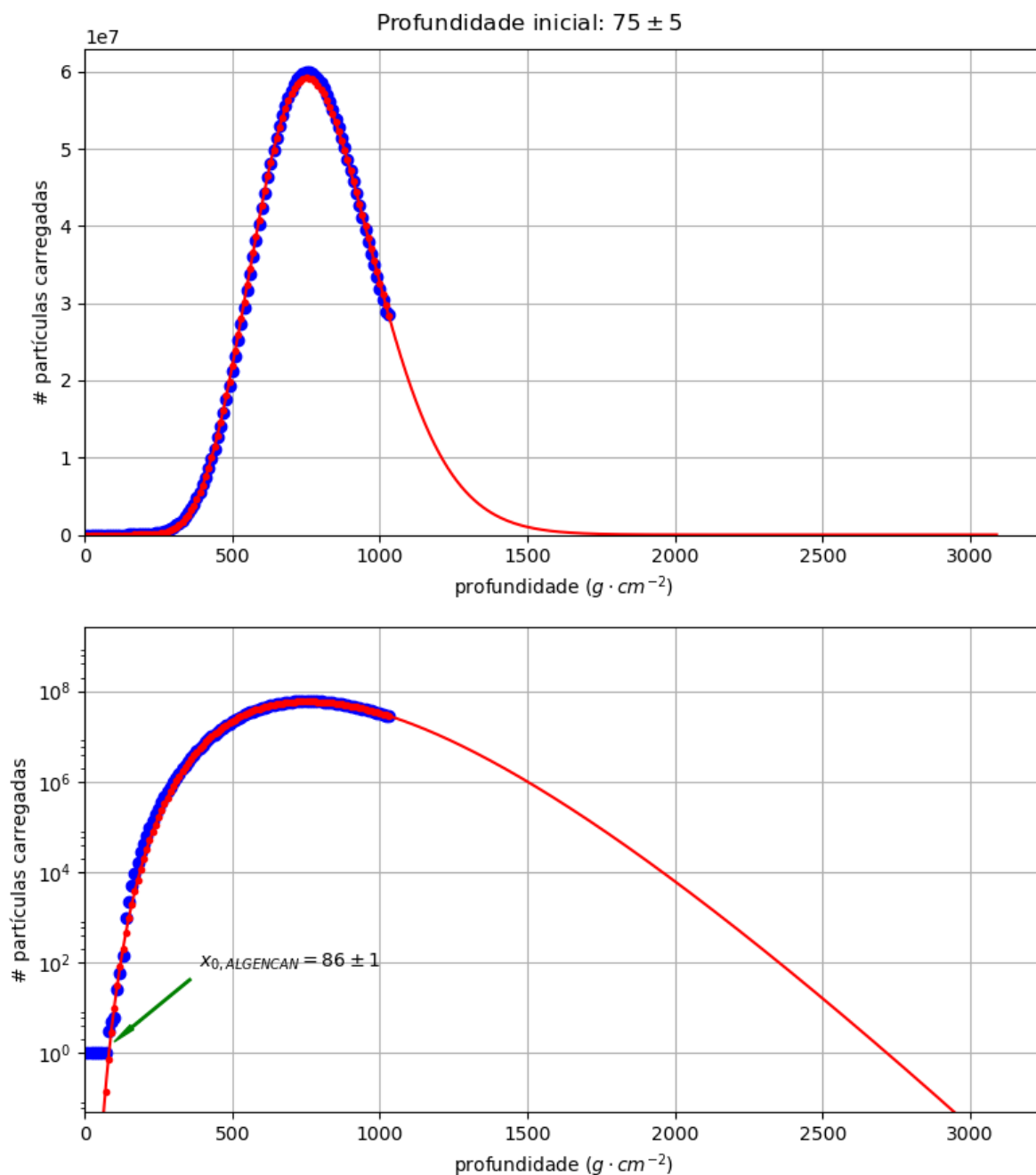


Figura 4.11: Resultado da aplicação do MEM em dados de colisão primária comparativamente profunda, com simulação da cascata de partículas até uma profundidade de aproximadamente $1000g \cdot cm^{-2}$, considerando a entropia da distribuição parametrizada até aproximadamente $3000g \cdot cm^{-2}$, com as condições de contorno de média linear, média logarítmica, e variância dos dados. Os dados são apresentados em azul, e a distribuição obtida é representada pela curva vermelha, com pontos nas posições correspondentes às quais há a presença de dados.

Na seção 4.6 é feita uma avaliação da determinação da profundidade primária.

4.6 Determinação da profundidade da colisão primária

Na figura 4.12 são apresentados os resultados da parametrização para obtenção da profundidade primária de diversos chuveiros. Como os dados foram simulados, a informação sobre a profundidade inicial do chuveiro está disponível; pode-se comparar esta informação, apresentada no eixo horizontal, com o parâmetro obtido de profundidade primária pelos diferentes métodos, apresentado no eixo vertical. Assim a linha verde representa o resultado ideal, em que as profundidades obtidas correspondem exatamente aos valores reais.

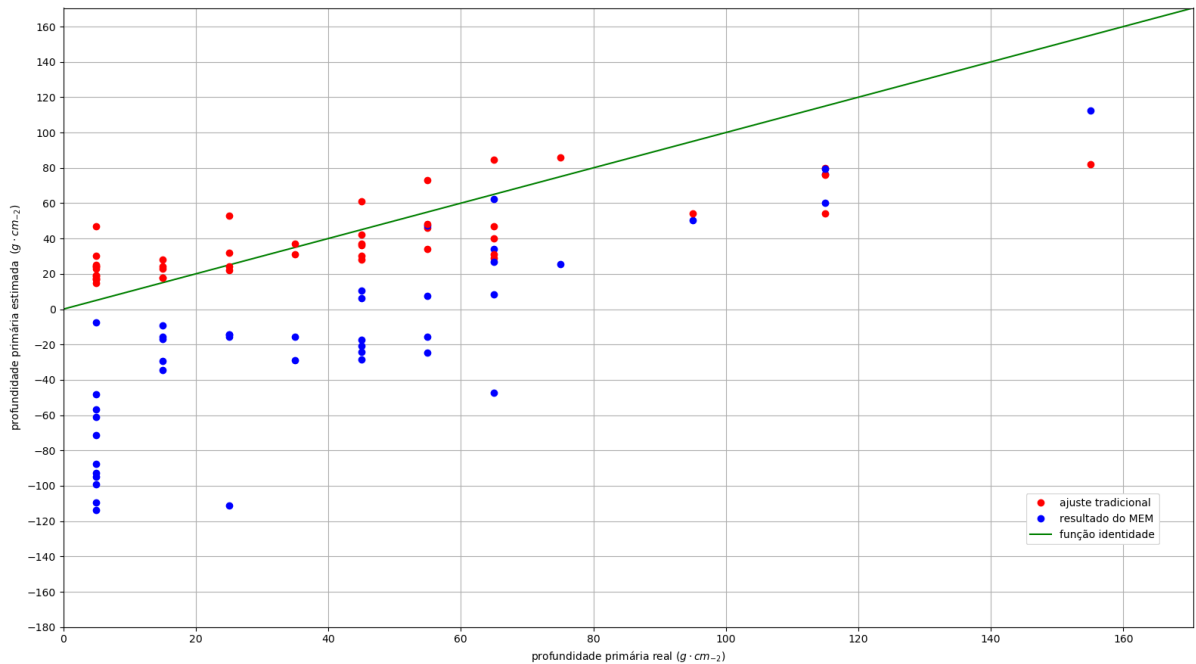


Figura 4.12: Resultado da parametrização da profundidade primária para diversos chuveiros atmosféricos. O eixo horizontal corresponde à profundidade primária real e o eixo vertical corresponde à profundidade primária obtida pelas formas de parametrização da distribuição comparadas. Os pontos azuis representam o resultado do ajuste de curva tradicional da função de distribuição de Gaisser-Hillas realizado automaticamente pelo CORSIKA, os pontos vermelhos os resultados obtidos pela aplicação do MEM com as técnicas e condições de contorno descritas neste capítulo, e a linha verde representa a região “ideal”, em que as profundidades primárias obtidas corresponderiam perfeitamente àsquelas dos dados simulados.

Percebe-se que o ajuste de curvas, apresentado com os pontos azuis, não só produz resultados distantes da reta ideal, como também produz valores negativos para a profundidade inicial. Isso não tem sentido físico para chuveiros atmosféricos, pois seria equivalente à ocorrência de uma colisão da partícula com os gases da atmosfera antes de a partícula ter entrado na atmosfera.

Os pontos vermelhos correspondem aos parâmetros obtidos pelo uso do método da

entropia máxima, cujos resultados ficam claramente muito mais próximos aos valores desejados. Dessa forma, o uso de dados vindos de simulações permite avaliar que o MEM produz resultados muito mais próximos dos desejados para a determinação da profundidade primária, sendo um método eficiente para a obtenção não só deste, como também os outros parâmetros da distribuição.

A natureza do método de não só atender as condições de contorno dos dados da distribuição, mas também levar em conta as regiões de dados indisponíveis e minimizar o viés da curva obtida ao buscar o maior grau de incerteza, permitem também um ajuste muito apropriado para a cauda da distribuição. Sua caracterização é considerada uma importante análise dos chuviros atmosféricos; a ausência de dados do fim da cauda (devido à perda do perfil a partir da profundidade da superfície do planeta) torna mais evidente quão adequado este método é para estudos realizados na área.

5. CONCLUSÕES E PERSPECTIVAS

Foram aplicados conceitos de teoria de informação a um fenômeno da física de raios cósmicos, que é o desenvolvimento de chuviros extensos na atmosfera. Realizada a descrição dos chuviros no capítulo 2 e no capítulo 3, abordou-se o método da entropia máxima, cuja aplicação a chuviros atmosféricos já constitui uma abordagem alternativa àquela tradicional. O próximo passo foi encontrar uma metodologia de cálculo numérico que contemplasse a visão entrópica introduzida na discussão do perfil longitudinal. O método dos multiplicadores de Lagrange foi escolhido para isso, sobre qual foram descritos os princípios gerais. A otimização numérica seguiu o programa ALGENCAN, que emprega o método do Lagrangeano Aumentado.

O procedimento aqui adotado apresentou bons resultados não só na parametrização da profundidade primária, mas mostrou-se muito apropriado na caracterização da cauda do perfil longitudinal dos chuviros atmosféricos, tanto no caso em que todo o perfil é conhecido quando no caso em que ele é truncado. O método aqui desenvolvido pode ser aplicado em fenômenos de difusão de partículas em forma de cascatas, com criação e “aniquilação” de partículas. Portanto, pode tornar-se uma ferramenta útil em calorimetria, que é um grande tópico de física experimental, não restrito ao estudo de raios cósmicos.

Uma grande vantagem do uso do método da entropia máxima é a facilidade em aplicá-lo à entropia total de um sistema complexo do qual são obtidas medidas experimentais diferentes; o exemplo para chuviros atmosféricos é a medida do perfil longitudinal por telescópios de fluorescência e o perfil lateral por tanques de detecção Cherenkov. A partir de uma parametrização da distribuição volumétrica da cascata de partículas, é possível calcular as suas projeções longitudinal e lateral, tal que o método da entropia máxima pode ser imediatamente aplicado aos perfis conhecidos para que seja obtido o perfil desconhecido.

As considerações sobre a teoria da informação e sobre a entropia máxima não estão ligadas ao ajuste tradicional do perfil de chuviros extensos ou à curva de Gaisser-Hillas original. Elas estão ligadas ao fenômeno de difusão de partículas, principalmente quando se tem uma multiplicação em cascata, com produção de novas partículas e aniquilação subsequente.

A utilização de dados de simulação e dados experimentais do Observatório Pierre Auger e a exploração das possibilidades do estudo da distribuição volumétrica de partículas na atmosfera são perspectivas de desdobramentos do trabalho aqui apresentado.

REFERÊNCIAS

1. Jaynes, E. T. Information Theory and Statistical Mechanics. *The Physical Review* **106**, 620–630 (1957).
2. Jaynes, E. T. Information Theory and Statistical Mechanics II. *The Physical Review* **108**, 171–190 (1957).
3. Heck, D., Knapp, J., Capdevielle, J. N., Schatz, G. & Thouw, T. *CORSIKA: A Monte Carlo Code to Simulate Extensive Air Showers* rel. téc. FZKA 6019 (Forschungszentrum Karlsruhe GmbH, 1998).
4. Beatty, J. J. & Westerhoff, S. The Highest-Energy Cosmic Rays. *Annual Review of Nuclear and Particle Science* **59**, 319–345 (2009).
5. Olinto, A. V. The origin of ultra-high energy cosmic ray: New physics and astrophysics. *Nuclear Physics B - Proceedings Supplements* **97**, 66–67 (2001).
6. The Pierre Auger Collaboration. The Pierre Auger Cosmic Ray Observatory. *Nuclear Instruments and Methods in Physics Research Section A* **798**, 172–213 (2015).
7. Shannon, C. E. A Mathematical Theory of Communication. *The Bell Systems Technical Journal* **27**, 379–423 (1948).
8. Lagrange, J. L. *Mécanique Analytique* (1788).
9. Birgin, E. G. & Martinez, J. M. *Practical Augmented Lagrangian Methods for Constrained Optimization* ISBN: 161197335X, 9781611973358 (SIAM, 2014).
10. Nagano, M. Search for the end of the energy spectrum of primary cosmic rays. *New Journal of Physics* **11** (2009).
11. Gaisser, T. K. & Stanev, T. Particle astrophysics and high-energy cosmic rays. *Physics Letters B* **667**, 254 (2008).
12. *Proceedings, 34th International Cosmic Ray Conference (ICRC 2015)* **ICRC2015** (2015). <<http://pos.sissa.it/cgi-bin/reader/conf.cgi?confid=236>>.
13. Penzias, A. A. & Wilson, R. W. A Measurement of Excess Antenna Temperature at 4080 Mc/s. *Astrophysical Journal* **142**, 419–421 (1965).
14. Greisen, K. End to the Cosmic-Ray Spectrum? *Physica Review Letters* **16**, 748–750 (1966).

15. Zatsepin, G. T. & Kuz'min, V. A. Upper Limit of the Spectrum of Cosmic Rays. *Journal of Experimental and Theoretical Physics Letters* **4**, 78–80 (1966).
16. The University of Adelaide. *Ultra High Energy Cosmic Rays (UHECR)* <<http://www.physics.adelaide.edu.au/astrophysics/hires/uhecr.html>> (2005).
17. Fretter, W. B. *Proceedings of the Echo Lake Symposium*, 39 (1949).
18. Bellido, J. & The Pierre Auger Collaboration. Mass Composition Studies of the Highest Energy Cosmic Rays. *Proceedings for the XXth Rencontres de Blois* (2008).
19. The Pierre Auger Collaboration. Testing Hadronic Interactions at Ultrahigh Energies with Air Showers Measured by the Pierre Auger Observatory. *Physical Review Letters* **117** (2016).
20. *Pinpointing the origin of ultra-high energy cosmic rays* <http://auger.org/education/Auger_Education/origincosmicrays.html>.
21. *A Hybrid Detector* <<https://www.auger.org/index.php/observatory/auger-hybrid-detector>>.
22. Alex F. Bielajew, Hirayama, H., Namito, Y., Nelson, W. R. & Rogers, D. W. O. *EGS4 Source Code Availability and Distribution* <http://rcwww.kek.jp/research/egs/egs4_source.html> (1997).
23. Ostapchenko, S. QGSJET-II: towards reliable description of very high energy hadronic interactions. *Nuclear Physics B - Proceedings Supplements* **151**, 143–146.
24. Fletcher, R. S., Gaisser, T. K., Lipari, P. & Stanev, T. SIBYLL: An event generator for simulation of high energy cosmic ray cascades. *Physical Review D* **50** (9 2006).
25. Pierog, T., Karpenko, I., Katzy, J. M., Yatsenko, E. & Werner, K. EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider. *Physical review / C* **92**, 034906. ISSN: 1089-490X (2015).
26. Jaynes, E. T. Gibbs vs Boltzmann Entropies. *American Journal of Physics* **33**, 391–398 (1965).
27. Gaisser, T. K. & Hillas, A. M. Reliability of the method of constant intensity cuts for reconstructing the average development of vertical showers. *International Cosmic Ray Conference* **8** (1977).
28. Hestenes, M. R. Multiplier and gradient methods. *Journal of Optimization Theory and Applications* **4**, 303–320. ISSN: 1573-2878 (1969).

29. Polyak, R. Modified barrier functions (theory and methods). *Mathematical Programming* **54**, 177–222. ISSN: 1436-4646 (1992).